# EDUCATION

A MONTHLY MAGAZINE

DEVOTED TO THE SCIENCE, ART, PHILOSOPHY AND LITERATURE OF EDUCATION

HERBERT BLAIR, EDITOR-IN-CHIEF
PROFESSOR OF SCHOOL ADMINISTRATION, BOSTON UNIVERSITY

VOL. LIII MARCH, 1933 No. 7

## MEASUREMENT NUMBER

DR. ERNEST W. TIEGS, EDITOR
DEAN OF UNIVERSITY COLLEGE AND PROFESSOR OF EDUCATION
UNIVERSITY OF SOUTHERN CALIFORNIA

## CONTENTS

| | | PAGE |
|---|---|---|
| Measurement and Progress in Education | *Ernest W. Tiegs* | 385 |
| Superior Scholastic Aptitude and Success in College. | *Frank C. Touton and Betty Trier Berry* | 387 |
| Measuring the Results of Religious Education | *Roy G. Bose* | 393 |
| Measurement in Educational Finance | *Willis W. Clark* | 399 |
| The Barrier (Poem) | *Rachel L. Dithridge* | 402 |
| Measurement in Determining the Personality and Behavior of the College Cribber | *William G. Campbell* | 403 |
| Dangers and Values in Teacher-Made Tests | *D. Welty Lefever* | 409 |
| How to Study for Objective Tests | *C. C. Crawford* | 413 |
| Some Results of the Testing Movement | *J. Harold Williams* | 417 |
| The Effects of Rest Periods on Progress in Learning in Typewriting. | *A. S. Raubenheimer and S. H. Davidson* | 421 |
| Needed Developments in Measurement for Secondary Schools | *J. Murray Lee* | 424 |
| Education in 1940 | *Frank C. Touton* | 428 |
| The Function of Examinations | *Mark E. Hutchinson* | 436 |
| Breaking the Foreign Language Goose Step | *Walter Vincent Kaulfers* | 440 |
| Your Schools and Mine | | 444 |
| Department of Book Reviews | | 446 |
| New Educational Books | | 448 |

Indexed in the Readers Guide and in the Education Index

**Published by THE PALMER COMPANY**
**120 Boylston Street Boston, Massachusetts**

Raymond P. Palmer, Business Manager

Published monthly except July and August. Subscription price $4.00 per year. 40 cents per copy.
Entered as second class matter at the Post Office, Boston, Mass., under the Act of March 3, 1879

## NEXT MONTH — PHYSICAL EDUCATION NUMBER

**Edited by DR. FREDERICK RAND ROGERS**

Dean of Student Health and Physical Education, Boston University

THE RELATION OF PHYSICAL EDUCATION TO THE HEALTH AIM.
By Frederick Rand Rogers.

BASIC PRINCIPLES AND PRACTICES OF PHYSICAL EDUCATION.
By Carl L. Schrader, Massachusetts State Supervisor of Physical Education.

HEALTH INSTRUCTION THROUGH THE PHYSICAL EDUCATION TEACHER.
By Dr. D. Oberteuffer, Professor of Physical Education, Ohio State University.

THE ADAPTATION OF PHYSICAL EDUCATION TO INDIVIDUAL NEEDS.
By Charles D. Glanque, Director of Physical Activities for Men, Boston University.

CORRECTIVE AND DEVELOPMENTAL ACTIVITIES FOR GIRLS.
By Eleanor Dobbins, Supervisor of Physical Activities for Women, Boston University.

FUNDAMENTAL DIFFERENCES BETWEEN BOYS AND GIRLS AS THEY AFFECT THE GIRLS' PROGRAM OF PHYSICAL EDUCATION.
By Mabel Lee, Professor of Physical Education, University of Nebraska.

THE FUTURE OF COACHING IN SECONDARY SCHOOLS.
By Dr. John M. Harmon, Director of Gymnastics and Team Coach, Boston University.

SELF-EDUCATION THROUGH PHYSICAL ACTIVITIES.
By Laurence S. Hill, Director of Physical Education, Ithaca College.

THE PLACE OF PLAY AND RECREATION IN THE COMMUNITY.
By Ernst Hermann, Director of the Sargent School, Boston University.

PHYSICAL EDUCATION AT LYNBROOK, NEW YORK.
By James J. Carter, Director.

PHYSICAL EDUCATION PROCEDURES IN QUINCY, MASSACHUSETTS.
By William H. Whiting, Supervisor.

**TEST** YOUR ART CLASSES BY MEANS OF THE

**Meier-Seashore Art Judgment Test**

Are your public school art courses reaching the ones who will profit most from the training?

No teacher of art should be without this useful device for locating persons who have outstanding artistic ability

*For use in Grade Seven through College.*

**PRICES:**

**TEST BOOK**, $0.90 per copy; five copies, $4.00; ten copies for $7.50; $0.70 per copy in lots of 35 or more; transportation extra.

**EXAMINER'S MANUAL**, ten cents, postage extra.

**TEST RECORD SHEETS** (one copy for each student tested), $2.00 per hundred, postage extra.

**SAMPLE UNIT SET**: Test Book, Test Record Sheet, Examiner's Manual, $1.10 postpaid.

**Bureau of Educational Research and Service**
Extension Division
UNIVERSITY OF IOWA, IOWA CITY

**OBJECTIVE TESTS IN AMERICAN LITERATURE**
*By* HAZEL CALL

Price: $2.00 per hundred; Specimen set, 25 cents

1. Literature in General
2. The Colonial Period
3. The Revolutionary Period
4. The First National Period
5. The Period of Conflict
6. All-American Period
7. Present Day Period

**THE PALMER COMPANY**
120 Boylston Street, Boston, Mass.

**CHARACTER *and* PERSONALITY TESTS**

Our latest series of tests developed in connection with the Character Education Inquiry under the leadership of Professor Hugh Hartshorne.

*Twenty-three tests of*

*Moral Knowledge*
*Opinion*
*Conduct*
*Reputation and*
*Background*

Write today for a free descriptive circular of these tests.

**ASSOCIATION PRESS**

**347 Madison Avenue** **New York**

# EDUCATION

Vol. LIII MARCH, 1933 No. 7

## MEASUREMENT AND PROGRESS IN EDUCATION

ERNEST W. TIEGS
UNIVERSITY OF SOUTHERN CALIFORNIA

ONE of our contemporary philosophers has said that in the ultimate analysis our troubles are due to dogma and deduction. We find no new truths because we take some venerable but questionable proposition as an indubiable starting point and never think of putting this assumption itself to the test of observation and experiment. For example, there once arose the belief that night air is bad; and come what might, education, ridicule, the pronouncements of medical science, and all other evidence to the contrary notwithstanding, many good folk of the past generation went to their graves with the firm conviction that night air is dangerous.

The field of education is not without its superfluous burden of dogma and superstition. Teachers on every level from the kindergarten to the graduate school still believe that they can adequately measure achievement or intelligence by looking at a student or hearing him speak; others have heard of intelligence tests and are under the impression that they are dangerous; still others believe that the passing mark is a simple, stable standard, in defense of which they would give their professional lives, rather than a fiction of the imagination. Still others believe that the traditional examination is a measurement instrument. Rating schemes without number and a host of impossible, amateurish devices flourish like the green bay tree in a profession which has the opportunity, the motivation, and the facilities for doing its work efficiently, effectively, and scientifically.

The task of applying the resources of measurement to practical problems in education has been complicated in the past, not merely by the dearth of experimental studies in many fields, which has forced teachers to pursue their activities unguided, or shall we say unhampered by facts; too often members of the profession have operated in distinct hostility to well-known facts. Before they were well launched in the profession to become their life's work, they have developed certain mind sets. They have hugged their illusions; they have feared changes in ideals, attitudes, and practices which would disturb their naive beliefs, to which they have clung as to a religion. Those to whom this description applies have soon lost their contact with the world of educational reality, and they have become strangers in a strange land.

We as teachers have agreed that it

is our objective to do more and better teaching, in order that the lot of children may be improved; and then we have done all in our power to prevent any change in the status quo, because it always involved additional learning and adjustment.

But the profession moved on; research revealed new truths; new devices and procedures were invented to apply these truths; but the thinking and educational procedures of only a very small part of the profession changed with needs and with the times.

Thus the world at large with its troubles and problems had its counterpart in the field of education — a relatively small group of leaders and a great body of more or less unthinking followers. The minds of the rank and file of the profession were cluttered with the dead husks of past reality which had long been discarded by those who had gone on learning and thinking.

During the past ten to fifteen years, a new spirit has permeated the profession. Workers in education have become aware of the fact that with the spread of democratic concepts in educational organization and practice, with greater freedom and greater responsibility for teachers, it is essential that they should receive a more thorough, scientific training and a wider knowledge and more profound understanding of their problems in order that this responsibility may be adequately discharged.

In his volume on "Making the Most of High School," Dr. Clyde Hill of the Yale School of Education, comments on the importance of mathematics in ancient civilization and the crucial part which it has played in the development of modern civilization, in the sciences, the professions, and in every-day life. Tests and measurements are simply the mathematics of Education. They are simply the means of substituting more accurate, reliable, and valid information to serve with other concomitants, as the rational basis for educational activities.

It is no longer necessary to guess at the extent of certain learning difficulties. They can be diagnosed. We need no longer merely argue about motivation; the effects of certain types of motivation may be tested in the same way that classification, ability grouping, guidance, marking and promotion, and other classroom activities may be studied,—in the light of objective evidence. The evaluation of textbooks and instructional procedures may be approached in a spirit which eliminates to a large extent the influence of personal prejudice or political pressure. The evaluation of teaching efficiency may proceed with a clear understanding of the pitfalls which in the past have attended it; however undesirable the implication, there is no question but that the average rating plan is constantly in danger of having much in common with discredited plans for character rating.

Better quantitative procedures are as inevitable in education as they have been in other fields. It is just a question as to how soon teachers receive adequate training for this phase of their responsibility. In spite of tradition, superstition, inertia, and misunderstanding, which constitute constant resistance to change, increasingly accurate knowledge has conditioned prog-

gress in other fields and gives great promise of unusual significance in education. The immediate task is to make teachers aware of the resources of measurement, and to provide them with the minimum working tools not only in expediting and improving classroom procedures, but in improving all techniques which are related to the educational process. Whether those techniques be related to improved teaching, improved materials of instruction, improved supervision, or improved administration, there is hardly a problem in which the classroom teacher does not have a vital interest. A training in measurement will serve not only as the means to progress, but as a method of evaluating whatever progress is made.

# SUPERIOR SCHOLASTIC APTITUDE AND SUCCESS IN COLLEGE

FRANK C. TOUTON
VICE-PRESIDENT AND DIRECTOR OF EDUCATIONAL PROGRAM
AND
BETTY TRIER BERRY
ASSISTANT PROFESSOR OF ORIENTATION
UNIVERSITY OF SOUTHERN CALIFORNIA

*The authors point out that the very superior students, as indicated by scholastic aptitude tests, are not as likely to make a superior record in college as are those with less ability. Is this the fault of the student, of the instructor, of the college curricula, or of the elementary and secondary school training?*

*Introduction.*

MANY investigations have been made and reported on in the field of prognosis of college marks. The relation between scholastic ability and academic performance is in general disappointingly low due, it seems, to two chief causes, namely, the unreliability of instructor's marks and the distractions of college life. On the former (unreliability of marks) no comment will be made other than to say that reliability of marks seems to vary with the objectivity, validity, and reliability of the tests used in measuring scholastic achievement. At this time more consideration will be given to factors which seem to influence academic success. Our attention will be focused on the levels of achievement attained by college students of very superior and superior scholastic aptitude. The single measure of aptitude used in cases reported here is the Thorndike Intelligence Examination for High School Graduates. The group considered is limited by the fact that this examination, though required of freshman entrants, is not taken by students admitted to sophomore and higher standing. The cases reported on will be in each instance less than the total group of entrants and will be selected from the higher levels of scholastic aptitude as measured by the Thorndike Examination. With the data available, an attempt will be made to seek an answer to the question: "Do very superior and superior college students achieve the scholastic success predicted by their differential mental capacities?"

*In the Experimental Entrance Group.* As has been announced in the

public press, the University of Southern California has now in progress a three-year plan, during which time from seventy to one hundred students will be admitted each year on a basis of entrance somewhat different from that used for admission to regular standing. Applicants who are accepted for this group are all graduates from accredited secondary schools and under twenty-one years of age. They have less than the required number of entrance units of recommending quality; they all have had the usual secondary school training and are recommended as to character and judged worthy of an opportunity to attend college for at least one year.

The final selection of members for this group is made by a committee which examines the entrance test records made by all applicants in a series of eight different tests. Students who are admitted thus are expected to achieve a scholarship average of 1.0 (C) or higher in the work of a full year in order to remain in the university. If they succeed in doing so, they become sophomores; if they do not achieve a scholarship average of 1.0, they are dropped from the university and may not be readmitted until they meet in full the regular entrance requirements.

During the past year a critical study has been made of the equipment, methods of study, and progress of this group.

In the first semester of the experiment, it was found that five members of this group were placed in scholastic aptitude among the highest seven per cent of the students regularly admitted to the university during the preceding three years. We would expect these five students to achieve excellent scholarship averages in their course work. Instead, four out of this group of five students failed during their first semester to make a scholarship average of 1.0. One failed almost completely and withdrew at the end of the first semester. Of the four who remained through the second semester, two failed in the work of the full year; of the other two, one barely passed. Thus there is only one of the five whose performance is consistent with his brilliant promise, and who has achieved a superior academic record.

When we consider the members of this group who were rated superior, though not very superior, we find a different result. Of the sixteen members of this group who are rated superior (B) in scholastic aptitude, thirteen (over 80 per cent) earned a satisfactory scholarship average during the first semester. One of these thirteen did not return for the second semester, while the three whose records for the first semester were unsatisfactory did return. During the second semester, then, fifteen students in the experimental entrance group were of superior scholastic aptitude. All but one of these finished the year with a scholastic average over both semesters of 1.0 or better.

It appears from our records of this group admitted and taught as experimental entrance students, those rating conspicuously high in scholastic aptitude were not as a group as consistently successful in their college work as were those members of the group who rated high in scholastic aptitude yet lower than the highest seven per cent. Conspicuously high aptitude records predict college success with less

certainty than do merely superior aptitude records. Superior performance in this group goes more often with superior than with very superior aptitude records.

*Records of regularly admitted students.* Since it appeared that superior performance went more certainly with superior aptitude than with very superior aptitude in the case of students admitted to the experimental entrance group, the question was raised as to whether this relation holds in the case of students admitted regularly. For this purpose the entering class of 1931 was chosen for study. There were 49 members of this class who were rated very superior (A) in scholastic aptitude. Of this group of 49 students, 32 (64 per cent) rated C or higher in their academic work of the first semester, while 17 (36 per cent) made unsatisfactory records, in many cases very unsatisfactory.

Some 161 students in this class were rated as just superior in scholastic aptitude as shown by the test. Of this superior group of 161 students, 132 (82 per cent) made satisfactory scholarship records during their first semester, while 29 students (18 per cent) did not.

The second semester records merely repeat the same story that is told by the first semester. Here 41 students of very superior scholastic aptitude completed the work of the second semester; 25 (61 per cent) did satisfactory academic work, while 16 (39 per cent) did not. For this second semester the group rated superior in scholarship numbered 141, of whom 114 (81 per cent) did satisfactory academic work, while 27 (19 per cent) did not. It appears that about two out of every three very superior and four out of every five of the merely superior students in scholastic aptitude achieve satisfactory academic records.

Even though the students rated superior (B) in scholastic aptitude are more certain to succeed as individuals than are the students rated very superior (A) in scholastic aptitude, yet the average achievement of the very superior group is higher than that of the superior group, for those of the very superior group who do succeed achieve very superior records.

It is recognized that the observations made above are drawn from a study of relatively few cases.

An explanation seems again to come from the close-up study of the members of the experimental entrance group. For these students, the records show that practically all of the very superior students are placed either above the ninety percentile or below the ten percentile point of the normal group in personality traits as announced by the Bernreuter Personality Inventory. Thus the very superior members of the Experimental Entrance Group show marked personality quirks and eccentricities. It seems reasonable to assume that, in general, students with marked scholastic aptitude may possess such personality peculiarities as will probably stand in the way of their success. Such students dislike routine tasks and require teachers of unusual ability and resourcefulness.

*Scholastic aptitude and marked achievement for the entire course.* The June, 1932, commencement program shows the names of 49 graduates who achieved *magna cum laude* or *cum*

*laude* mention, and for whom the university records show scholastic aptitude scores taken during the freshman year.

Of the 13 students who achieved the *magna cum laude* recognition and for whom scholastic aptitude records were on file, 5 (37 per cent) were in the very superior, 7 (55 per cent) in the superior, and 1 (8 per cent) in the average scholastic aptitude group. Of the 36 students who achieved the *cum laude* recognition and for whom aptitude records were on file, 4 (11 per cent) were in the very superior and 4 (11 per cent) in the average, while 28 (77 per cent) were in the superior aptitude group.

Here we see that the superior rather than the very superior scholastic aptitude students achieve a higher percentage of both the *magna cum laude* and *cum laude* graduation scholarship recognitions. The highest places on the commencement program seem to have been reserved for the most part for students who showed only superior scholastic aptitude at entrance, a few being left for very superior students in scholastic aptitude, and an occasional place is held for the college entrant of average scholastic aptitude.

Reference to another commencement program confirms the above findings. Here we have scholastic aptitude scores for 6 of the 1931 graduates with *magna cum laude* and for 24 graduates with *cum laude*. One *magna cum laude* graduate (18 per cent) rated very superior in scholastic ability at entrance and 5 (82 per cent) superior in scholastic ability. Of the 24 *cum laude* graduates for whom records were available, 4 (11 per cent) were rated very superior in scholastic ability at college entrance, 18 (67 per cent) superior, and 2 (8 per cent) were of average scholastic ability.

*Scholastic aptitude and membership in honor societies.* Two nationally known scholastic honor societies elect members each year from among graduates who earn academic records placing them in the highest ten per cent of the graduating class. Since certain of the members elected to these societies entered the university as freshmen, it is possible to examine their scholastic aptitude records and note the relationship which exists between very superior scholastic aptitude records and election into scholastic honor societies.

Of the group elected in 1932 to Phi Beta Kappa, a national Liberal Arts scholarship society, 19 seniors had scholastic aptitude records on file. Of the 19 seniors, only 4 (21 per cent) had very superior scholastic aptitude records, while 15 (79 per cent) had superior or average records. In June, 1931, 11 seniors whose scholastic aptitude records were on file were honored by Phi Beta Kappa election. Of these 11 students, 3 had very superior scholastic aptitude records while 8 had records showing only superior or average scholastic aptitude.

For Phi Kappa Phi, a national scholarship society which draws its members from the several schools and colleges of the university, the story is much the same. In 1931, there were 39 members elected for whom scholastic aptitude records were available for this study. Of this group, 11 (28 per cent) rated very superior in scholastic aptitude at college entrance, while 28 (72 per cent) rated superior or average in aptitude. In June of 1932, there were 50 new members elected

whose aptitude records are known, only 11 (22 per cent) of whom entered with very superior scholastic aptitude records.

The above facts indicate that membership in national scholarship societies is achieved not alone by the students who show very superior scholastic aptitude at college entrance, for the membership lists for two consecutive years show the names of from three to four times as many superior and average aptitude students as of students of very superior scholastic aptitude.

*Correlation between aptitude and achievement.* When a scatter diagram is made showing the relation between scholastic aptitude and academic achievement for entrants of September, 1931, we find a Pearson correlation of $0.32 \pm 0.02$, a finding which is in agreement with the results of previous studies. The group of 1931 freshman entrants may be taken as a typical University of Southern California entrance group. The relationship shows two distinct trends: in one there is notable consistence of performance with aptitude—high performance with superior aptitude and low performance with inferior aptitude—while the other indicates that there are a considerable number of students with very superior aptitude who show average or low achievement performance.

The scatter diagram shows this fact quite clearly. Most of the scores take their places in the diagram relatively near to the diagonal which runs from lower left to upper right of the field. A few records, however, form a sort of wedge, broad in the upper left-hand corner and narrowing toward the center of the field, where it is merged with and lost in the other records.

A study of this correlation seems to give support to the belief that among students of very superior scholastic aptitude, two distinct sets of performances may be expected. Some achieve superior mastery of conventional subject matter with the usual methods in class group situations, while a considerable part of these very superior aptitude students possess erratic traits and personality conflicts which make it difficult for them to adapt themselves to the usual school situations. They seem to fail to develop social group traits sufficiently well to succeed even approximately near the level of their capacities. In fact, many find themselves so hopelessly out of step with other students that they either drop out of college altogether or, if they remain in college, seek other major avenues of self-expression and continue on in course work with but mediocre achievement.

*Résumé.* An answer is sought to the question "Do students of superior rather than very superior scholastic ability achieve with greater certainty the superior academic records?"

Of the students admitted in the 1931 experimental entrance group at the University of Southern California, those rating conspicuously high in scholastic ability are not as a group as successful in their college work as are those members of the group who are rated as superior, yet less brilliant. For the group, very high aptitude records predict college success with less certainty than do merely superior records. With the experimental entrance group superior performance goes more

often with superior than with very superior scholastic aptitude records.

The study of the very superior and superior members of the 1931 group of entering freshmen at the university indicates that four-fifths of the superior students and only two-thirds of the very superior students in scholastic aptitude achieve satisfactory academic records. Personality quirks and personal eccentricities seem to impede the progress of many scholastically very superior students. These seem to adjust and adapt themselves less well to the usual school environment and procedures.

From a study of entrance scholastic aptitude records of students who are graduated *magna cum laude* and *cum laude* it appears that students of superior, rather than very superior, scholastic aptitude achieve the higher percentage of the conspicuous scholarship honors at graduation. Places of highest honor on the commencement program seem to be reserved for students who show superior rather than very superior scholastic aptitude at college entrance.

A study of the Phi Beta Kappa and Phi Kappa Phi national honor scholarship societies' graduation lists shows that membership in national honor scholarship societies is achieved not alone by the students who show very superior scholastic aptitude at entrance to the university, for the membership lists of these societies show from three to four times as many names of students of superior and average scholastic ability as appear for students of very superior scholastic aptitude.

An analysis of the factors which produce the seemingly uncertain relationship between scholastic ability and academic achievement for college students indicates two distinct trends: a major trend of positive relationship between superior ability and high achievement and another counter, yet marked, trend which indicates that many very superior students do mediocre or failing work while some students of mediocre talent do work of a surprisingly high grade.

The answer to the question stated at the outset seems to be: Students of superior, rather than very superior, scholastic ability seem to achieve with greater certainty the superior academic records.

# MEASURING THE RESULTS OF RELIGIOUS EDUCATION

Roy G. Bose

CALIFORNIA CHRISTIAN COLLEGE, LOS ANGELES

*Professor Bose makes a plea for a better appreciation of the value of tests as teaching instruments. Religious and character education are no exceptions to the general rule that only through measurement can we discover the relative effectiveness of our teaching methods.*

### *Relation of Schoolman to Religious Education.*

SHOULD a schoolman be interested in evaluating the results of religious education? My answer is emphatically in the affirmative. The child is a unit. His head cannot be sent to school, his body to the playground, and his soul to church. He takes his entire self with him in all his explorations of life. Therefore the process which educates him must be unified, and the one who supervises his education must be interested in that entire process. He should know what results are being obtained in other fields than his own.

I am not advocating that the schoolman should become responsible for the religious education of the child. There is a vast difference between an intelligent interest in this phase of a complete education, and using the public school as an agent in fostering it. Because of his position of educational leadership in the community, the schoolman should be able to furnish expert advice, and leadership, particularly in testing and statistically evaluating the results of religious instruction.

In some fifteen hundred communities in the United States at the present time, week-day schools of religious education are being conducted. Many of these are on "release-time" or periods given during the school day, when the child is released from the school to go to a place designated by the churches for religious instruction. How can the schoolman effectively evaluate such work without a knowledge of the methods and materials now available for doing so?

Other considerations demand that the attention of the schoolman be directed to the techniques developed by the leaders in the field of religious education. Character education, which is rapidly becoming the major concern of education, has found its greatest ally in the leadership of religious education. Hartshorne and May, leaders in the "Character Education Inquiry," G. H. Betts, of Northwestern University, W. C. Bower, with his book "Character through Creative Experience," Goodwin B. Watson, author of "Experimentation and Measurement in Religious Education," and a host of others, work effectively in both fields. Religious education is increasingly emphasizing character outcomes and personality development. The schoolman who would develop intelligently a program of character education in his school cannot ignore this leadership.

### *Religious Education and Character Education Defined.*

I have used, repeatedly, the terms religious education and character edu-

cation. They are easier to use than to define. For the purpose of this discussion, I will define religious education as teaching having a religious motivation and character education as education with the "cultural" and "citizenship" emphasis. Immediately you question the distinction, and ask if religious education does not issue in cultural values and good citizenship? Therein lies the present confusion. The outcomes are the same. The motivation is somewhat different.

Another question may be raised as to the distinction between these forms of the educative process and that which we commonly call "Education." The older knowledge-centered viewpoint made the learning of the tool-subjects the end of education. Religious education had no place, and character outcomes were only incidental. With our newer emphasis of citizenship, education and character education are becoming synonymous.

Those concerned with the educative process, therefore, must be interested in evaluating the results of the character development in all phases of life. It becomes the duty of the schoolman to know the techniques in the field of religious and character education as well as in his own.

*Test Materials and Methods in Religious and Character Education.*

We now come to the problem which occasioned the writing of this paper. It is, in brief, What materials and methods are available at the present time for measuring the results of religious and character education? The answer must be introduced by the following statement. There are no techniques available for religious education that are not equally open to any other phase of education. Religion has no unique forms, which cannot be measured by the methods now being evolved. That which exists in some form can be measured in some way. The task is but to find the right measuring rod. The difference between religious education and the other forms of education lies in the materials used, and the outcomes to be measured. The measurable quantities are knowledge and attitudes.

In the realm of knowledge outcomes, a number of tests have been evolved. A few of these are: Whitley's Biblical Knowledge Tests; Giles' Sunday School Examination A; Laycock's Test of Biblical Information; Union Test of Religious Ideas; Fah's Biblical Test, and Bose's Religious Concept Test. A more complete listing of tests of this character, together with the name of the publisher, price, purpose, validity and reliability, will be found in Watson's Experimentation and Measurement in Religious Education. Another source of supply is the Association Press, New York.

In the realm of attitudes, which is the most important emphasis in religious and character education, there is an even larger number of available tests. Such tests as Watson and Chassell's, Emotional History Record; Fernald's, Ethical Perception Test; Hart's, Test of Social Attitudes and Interests; Koh's, Ethical Discrimination Test; Union Test of Ethical Judgment; Chapman's, Test of Motives, and the Y. M. C. A. Character Growth Tests are representative in the field of measurement of attitudes. Many others will be found listed in the sources given above. These tests are usually difficult of administration, and

should be given only by some one trained in technical procedure.

*Construction of Tests.*

In addition to the available standardized tests, one who would measure the results of religious and character education must know how to construct his own tests. Most of the tests constructed in the local community or the local church school must necessarily be to measure Biblical or other information. Attitude tests are of such a difficult nature as to require a highly technical procedure, not available in the average local situation. It would be preferable to use standardized tests in cases where attitudes are envolved. The leaders and teachers of the average church school could be taught to construct the simple information tests, which aid materially in evaluating their work.

As to the types of objective tests possible of construction for the measuring of information, the following are the most common.

1. *True-False Question Tests.*

A statement is given, and the subject indicates by "true" or "false," or plus or minus or a "T" or an "F," whether he believes the statement to be true or false. The following are simple statements.

a. Peter was known as the beloved disciple. True. False.
b. The 23rd Psalm was written by Isaiah. True. False.

The above form is easy to score, is purely objective, and allows for extensive sampling in a limited time. There are certain limitations to the use of the true-false questions, mainly, the danger of implanting false ideas. The success of this method is conditioned by the care taken in the construction of the statements.

2. *Completion Question Tests.*

A statement is given from which certain significant words have been omitted. The following are sample statements.

a. The Gospel of ..... is the third book of the New Testament.
b. ..... was the eldest son of Isaac.

The one answering the question or completing the statement fills in the blank with the word or words he feels to be right. This form of test is easy of construction, but difficult to score, because in some cases a variety of answers may be possible. The number of blanks should be kept small, preferably one. Each blank should call for only a single idea or word. This type of test possesses the advantage of emphasis upon the right answer.

3. *Multiple-Choice Question Test.*

This type of question is made up of a statement, which may be completed by underlining one of several possible choices.

a. The Lord's Prayer pictures God as:
Judge, King, Avenger, Father.
b. The Parable of the Lost Coin teaches us that:
1. We should never lose money.
2. We should sweep the house often.
3. All people are careless.
4. Everyone is precious in the sight of God.

The advantages of this type of test is that it is objective, and easy to construct. A process of judgment is required in questions of opinion, etc. The limitations lie in the construction, wherein the possible choices are not clearly differentiated. There should

be at least four or five responses, each possessing some degree of plausibility.

Several other types of objective tests of easy construction could be given, such as the *Matching Tests,* where lists of answers are matched with itemized statements; the *Simple Recall* Tests, which are simplified completion tests. A complete analysis of these tests may be found in such a book as Ruch's The Objective or New Type Examination.

*Principles Underlying Test Construction and Selection.*

A few simple questions should be raised with those whom we would train in the construction and selection of tests in these fields. They are:

1. *What do I want to test?* The kind of information or the particular phase of character development should be decided upon. Also it should be remembered that all materials are not equally important, nor can a test be all inclusive. The selection of the important items according to some plan is essential.

2. *What method of testing will most adequately reveal the information or attitudes I am seeking to find?* With a choice of a score or more possible methods, a selection on the basis of end in view must be made.

3. *How can I be sure that the test measures that which it purposes to measure?* One must know something of its validity and reliability, and the method of its standardization. A knowledge of elementary statistics, thus, becomes essential to anyone working with tests, no matter what the field.

4. *Which tests will fulfill the above conditions and yet have an ease of administration, so that they may be used by others than experts?* The form of test, the adequacy of directions, and the objectivity of scoring must be considered in constructing or selecting tests. The tests should be of such a character that they could be administered by a fairly well-educated person, who has been given some training in this field.

5. *Is the test simple and easy to score?* The test which requires an elaborate and intricate system of scoring will discourage the average user.

6. *In what ways can the results of the test be made available in furthering personality growth?* The ultimate value of any test lies in its usefulness as a means of furthering character development. To merely gather information is vicious. Also the tests should not be made to say more than the contents justify.

*Necessary Statistical Measures.*

In order to understand the meaning of standardized test terminology and to interpret the results of the tests given, one must know a number of more or less technical terms. Certain simple procedures such as finding the mean, mode and median, etc., are essential to any understanding of the use of measures. Scores take on meaning as they are related to other scores. Johnnie's rating is understandable only in relation to the others of his class, etc.

One method of so relating scores is by *distribution,* either from lowest to highest, or by ranking highest as 1, next as 2, down the scale. If one would find out how the group reacted as a whole, a *frequency distribution* of the scores is necessary, which reveals how many persons made each score. When the distribution has been

determined, a number of things may be learned about the group. First, the *mean* or *average* score of the group may be found by adding the scores and dividing by the total number. This will give the score which is the average of the group. With this measure the group may be compared to other groups.

Other means of *central tendency* are the *mode* and *median.* The mode which connotes the popular or most frequent score, and the median which indicates the middle score of the group are both easily found and are good measures of comparison. The one who would interpret results should know, further, how to divide the distribution into *quartiles,* so that each quarter group might be studied. Other methods of interpretation such as the *average* and *standard deviation, correlation, reliability* and *validity,* etc., are necessary factors in understanding the results of tests.

*Use of Results.*

The information thus gathered and evaluated should be presented to the administrators of the church school, the teachers, and, in some cases, to the parents. The problem of the most effective method of presentation of results is bound to be raised. The written report is usually confusing in the multitude of its findings. A better method is that of *graphs,* which picture the results in high relief. Again it will be found necessary to teach some techniques of graph making to those who aid in presenting the results of testing.

In this brief article, I have sought to show the experimental approach can be and is being applied to the problem of measuring the results of religious and character education. The greatest lack in these fields is not adequate tests, but trained workers, and an appreciation of the value of tests. Therefore I have appealed to the schoolman to bring his training and leadership into this field. I have listed a few of the standardized tests in the field, and suggested a few methods of constructing and interpreting simple tests. As the article draws to a close, it would be well to emphasize a point of view that we must promote, namely, that anything that exists can be measured if instruments can be constructed to do so. Tests are merely objective ways of measuring. There will be found those who contend that religious attitudes, emotional reactions, etc., cannot be measured, because of their variable nature. The only answer is that they are being successfully measured. There is still much to be done in this field, as the measurement of personality factors is far more difficult than the measure of knowledge outcomes. If we could educate the church school leadership to take an experimental attitude, wherein they seek to evaluate the results of the work done in an objective way, we would have made considerable progress.

There are definite limitations to measurement as now evolved. The limitations are found in the inaccuracy of the tests, and the uncritical acceptance of results of tests. Too great reliance upon results obtained by the present group of tests is as unscientific as the rejection of all results. A middle ground of appreciation and sane use is the point of view to urge upon the ones concerned.

Some will argue that the final test

of religious and character instruction is not in the response to tests, but in life itself. To this I would heartily agree, but I would insist that I be privileged to evaluate that life objectively at different stages along its progress of development. We do not wait until the child has become an adult to find out his mental equipment, but measure his progress at various stages along the way. Let those in other fields be as wise. Many times the cry is raised that common sense judgment is enough, but common sense should tell us to use the most scientific methods in discovering values.

The reader has noted, I am sure, the great similarity between measurement in religious and character education, and that in the field of general education. One can readily see that the methods must be the same. There is not one scientific method for religion and another for the other phases of life. The points of view are the same; the materials and ends are different. Therefore, it is hoped that religious and character instruction and general instruction may find an increasingly common ground in the great task of developing worthy personalities.

In some phases of measurement, religious and character education have made greater progress than has general education. General education has been mainly concerned with discovering I. Q. and A. Q. ratings for the purpose of classifying pupils, while the other branches of education have concerned themselves with personality growth outcomes. The outstanding work of the "Character Education Inquiry" under the leadership of Hartshorne and May in this connection is a notable example of excellent scientific procedure in determining character outcomes.

The future success of all education lies in its united effort and its scientific attitude. Increasingly the common aim must be character development. All the forces in the community must be made to contribute to this end. Means of evaluating these forces must be evolved. In this the schoolman will play a large part.

# MEASUREMENT IN EDUCATIONAL FINANCE

WILLIS W. CLARK

DIRECTOR OF ADMINISTRATIVE RESEARCH
LOS ANGELES COUNTY SCHOOLS, LOS ANGELES, CALIFORNIA

*Comparison with someone else is the common basis of measurement. "How do our salaries, buildings, cost per pupil, costs for operation of the plant, etc., compare with other schools?" These are questions constantly asked of every school superintendent. Mr. Clark points out how a better system of measurement will enable the superintendent to give a more correct answer.*

THE financial basis for the support of public education is being given much serious attention at the present time. During prosperous years the educational program has been considered of sufficient basic importance that requests for support by the educational administrators have been approved without unusual public attention to the financial factors involved. At the present time, however, the marked reduction in annual income and sources of revenue has resulted in a very critical attitude concerning educational expenditures. The topic under discussion, therefore, is one which merits careful analysis and evaluation in order that effective techniques may be applied in estimating and evaluating the financial problems of school administration.

Some of the primary factors in considering problems of educational finance are

(1.) The definition of the measurable units in the field of educational finance and

(2.) The application of measurement techniques to the analysis and interpretation of financial data.

*Measurable Units.*

One of the most frequent sources of error in financial studies is the comparison of variable units. Frequently comparisons are made between groups of data which are not of a similar definition. For example, in school statistics it is a frequent error that teacher load determined by use of enrollment is compared with teacher load determined by use of average daily attendance. There is an expected difference of 5 per cent or more in the results obtained by the use of these two measures.

Some of the most frequently used units will illustrate difficulties in measurement of educational finance:

1. Cost per pupil in average daily attendance.
2. Teacher unit.
3. School district.
4. Tax rate.
5. Assessed valuation.
6. Expenditure classifications, such as general control, teachers' salaries, operation of plant, etc.

Cost per pupil in average daily attendance would appear to be a relatively stable unit as it represents pupil days in attendance throughout the school year. Many studies have failed to take into account the fact that average daily attendance is determined by dividing the total attendance throughout the school year by the number of days taught. In other words, the school district operating 190 days might have the same average daily attendance as another district operat-

ing 170 days. The additional twenty days taught will result in an increased cost per pupil although this is not shown by data ordinarily published in connection with cost per pupil. To make this illustration more definite the following illustration will be of interest:

| School | Days taught | Cost per pupil | Correction to 170 days |
|---|---|---|---|
| A | 195 | $106.50 | $92.85 |
| B | 170 | 102.25 | 102.25 |

For purposes of educational finance, a teacher unit may be defined as 35 pupils or a fraction thereof in average daily attendance.[1] This term in itself is a variable factor of considerable consequence in smaller districts. For example, a school of 30 pupils in average daily attendance has one teacher unit, whereas a school with 40 pupils in average daily attendance has two teacher units. It appears that this measure may be used satisfactorily only in larger school districts where the error involved will be a negligible factor.

Teacher load is a somewhat similar term and is variously defined. In some studies it will include total pupil contacts per week, including home rooms and study periods. Frequently teacher load for schools with 45 minute periods are directly compared with schools with teacher loads for 60 minute periods.

School districts are probably the most frequently used unit of measurement in educational finance. Expenditures and pupil costs are computed for annual reports on the basis of such district groups. A comparison of the expenditures of two districts should take into account both the size of the district, the length of the school year, and special conditions which may have a bearing on the districts under consideration, such as unusual transportation charges.

Many studies of educational finance are dealing with the matter of school costs in which the total school expenditures broken down into the customary headings of general control, teachers' salaries, other expenses of instruction, library, operation of plant, maintenance of plant, fixed charges, transportation, and auxiliary expenditures. Some of the difficulties attendant upon the comparisons of historical data, data for different districts and particularly for districts in different counties and states, are the classification of expenditure items under the general ledger headings. These frequently change year by year and depend to a considerable extent upon the interpretation of the auditors or bookkeepers who are posting the accounts.

At the present time tax rates are subjects of considerable discussion. Presumably the tax bill received by the property owner, which specifies a given rate per hundred dollars of assessed valuation, would appear to be definite, specific, and unquestionable. However, for purposes of research and comparative statistics, there are a variety of factors involved. Among these are the following: The total tax rate consists of a number of separate rates for various purposes. Even the total-for-schools rate, which is a part of the total tax rate in Los Angeles County, consists of eight items. These rates are as follows:

A. Elementary.
   1. General County Elementary.
   2. District Elementary.

1 California School Code, section 4, 770 ff.

3. Elementary Building.
4. Elementary Bonds.

B. Secondary.

5. General County High School.
6. District High.
7. High School Bonds.

C. Junior College.

8. Junior College Tuition.

For purposes of research, it is necessary that tax rates be broken down into the specific purpose for which they are intended, and any discussion of this particular matter is more or less invalid if the persons concerned do not have basic information. For example, an apparent increase in the cost for schools for a certain school district in which the rate for 1931 was $2.76 per $100 assessed valuation and for 1932 was $3.02, the increase, which was attributed to costs of schools, was really due to a change in the bond rate from .36 to .78 which more than accounted for the increase.

Another factor of unusual significance in the comparison of tax rates are assessed valuations. The basis of assessment varies from county to county and state to state. Even though the ratio of assessed valuation to true valuation is presumably constant in a given county, there are significant changes in valuations in given school districts from year to year. For example, a comparison of the assessed valuations of two extreme instances for the years 1931 and 1932 shows that in one instance the assessed valuation increased 55.7 per cent and in the other instance the assessed valuation decreased 49.0 per cent. The average (median) decrease for Los Angeles County for 1932 was 19.5 per cent. As the tax rate is determined by dividing the authorized request for revenue by the assessed valuation, it is obvious that the ratio of assessed valuation to true wealth and also the changes which are occurring from year to year are factors of consequence in making comparable measurements in this field.

At this point it should be mentioned that the customary unit of measurement in all financial matters, namely, the dollar, is itself a variable quantity in terms of its purchasing power. Commodity prices and index numbers in this respect indicate the necessity of evaluating our original data, ratios, and costs in terms of some basic criteria such as the purchasing power of the dollar at a given time, for example, the year 1913.

*Measurement Techniques.*

Statistical method provides measures of central tendency, variability, and relationship. Each of these measures is applicable to the study of problems of educational finance. In determining the trend or most characteristic situation, the mean, median, or mode is most frequently used in the analysis of the data under investigation. These measures are particularly useful in showing changes in time and significant differences in the comparison of groups.

Measures of variability are equally important. Variation is ordinarily shown by presenting data in tabular form or it may be indicated by giving range, percentile ranks, standard deviation or quartile deviation. In case large numbers of instances are under consideration, the latter measures are necessary for interpretation of data.

These measurement devices are applicable to the study of problems of educational finance. They will assist

in establishing standards or norms for various sizes and types of districts. For illustration, the following table presents data for the expenditure item "operation of plant" for three groups of school districts.

Such information based on an analysis of an adequate sampling is essential for establishing minimum standards and in providing a basis for expenditure control.

Despite the difficulties confronting research workers in securing educational finance data based on constant and clearly defined units, the imperative need of accurate and adequate information is evident. A considerable part of the administrative difficulties in financing a satisfactory program of education is due to a lack of information of sufficient validity to command respect. Undoubtedly, the numerous studies now in progress will result in the development of measures and standards which will give the much needed definition and objectivity to educational finance.

EXPENDITURE PER PUPIL IN 1931-32 FOR OPERATION OF PLANT

| | 2-5 Teacher Schools Districts | | 6 Teacher Schools to 599 A.D.A. Districts | |
|---|---|---|---|---|
| | Low* | Average* | Low | Average |
| Janitors and Engineers' Salaries | $6.67 | $10.45 | $5.84 | $7.19 |
| Supplies for Operation | 0.95 | 1.29 | 0.73 | 0.94 |
| Fuel and Heat | 1.02 | 1.20 | 0.62 | 0.87 |
| Light | 0.42 | 0.67 | 0.26 | 0.55 |
| Water | 0.00 | 0.00 | 0.24 | 0.43 |
| Miscellaneous | 1.90 | 0.89 | 0.00 | 1.00 |
| Total per pupil | 10.96 | 14.50 | 7.69 | 10.97 |
| Number of Districts | 28 | | 34 | |

* For purposes of this tabulation, "Low" is the 25 percentile district or that district which is located at the point which divides the lower one-fourth from the upper three-fourths of the districts in the group. The "Average" is the median or middle district. The data are for Los Angeles County School Districts.

---

## THE BARRIER

Usually I'm content
With this, my chosen task;
But sometimes, as today,
I feel the unspoken arrogance
Of Youth toward Age.

Their young eyes say,—
What their lips will not utter:
How can you enter
Into our surging life?
How can you understand
The impulses that govern
And the needs that urge?
Your many words
Are foolish in our ears.

And so I know
Between us looms,—
Dark and impenetrable,
The barrier of the years.

—Rachel L. Dithridge,
Hollis, Long Island.

# MEASUREMENT IN DETERMINING THE PERSONALITY AND BEHAVIOR OF THE COLLEGE CRIBBER

WILLIAM G. CAMPBELL

UNIVERSITY OF SOUTHERN CALIFORNIA

*Professor Campbell shows that cribbing is widely prevalent in large classes, due not only to the pressure exerted for high marks, but also to the ease with which cheating could be accomplished.*

DURING the spring of 1931, an extensive study of student honesty was conducted at one of the large state universities in the South. Since no previous large-scale study of the personality of the college cribber had been made, it was thought desirable to gather data regarding the personality traits of the subjects whose actions in various situations were being investigated. One of the studies in this field which has been reported was that by Brownell.[1] He analyzed the test ratings of thirty uncaught college cribbers, and found that fairly definite groups of traits were common to the majority. His numbers were small, however, and only a limited number of test ratings were available.

In order to insure a fair sampling of the undergraduate student body and some range in faculty-student relationships, six classes were chosen from the departments of Educational Psychology and History of Education. Although in no case was the composition of the class limited rigidly to students of a given stage of scholastic advancement, two of the classes were designed primarily for students at the freshman level, two for students at the sophomore level, and two for advanced and graduate students. The numbers of students of each sex in each college year may be seen in Table I. Although the numbers of seniors and graduates were not extremely large, there was a sufficient number at each of the other levels to insure reasonable reliability for the experimental findings.

TABLE I

RELATIVE COMPOSITION OF THE GROUP WITH RESPECT TO CLASS AND SEX

| Class | Men | Women | Total |
|---|---|---|---|
| Freshman | 23 | 38 | 61 |
| Sophomore | 3 | 36 | 39 |
| Junior | 12 | 33 | 45 |
| Senior | 3 | 13 | 16 |
| Graduate | 6 | 6 | 12 |
| Total | 47 | 126 | 173 |

*The spy system.* The appelation of "cribber" was given to those students who were observed cheating on one or more course examinations, or who, when quizzed regarding their actions, admitted having done so. The remainder of the students in the group were classed as "non-cribbers," despite the fact that under other circumstances they might have been found guilty.

The cheating during regular examinations was detected by advanced students who acted as spies. A pre-arranged order of seating placed these spies at vantage points, and the small number of subjects (from six to nine)

1 H. C. Brownell, "Mental test traits of college cribbers," School and Society, 27: 764, 1928.

assigned to each spy permitted careful observations of the actions of each subject. A student was not accused of cheating unless he used specially prepared notes, a notebook, or a textbook, or unless he asked another subject a question and received an answer which would be of aid to him. In this latter case, both question and answer had to be overheard by the spy in order to count against the student. Since none other than these clear-cut evidences of procuring information from illegitimate sources were scored against a student, it is reasonable to assume that the figures concerning the cheating which occurred during examinations are conservative. Looking on another student's paper, since the spy could not be certain that aid had been received, was not counted as cribbing. Many of the students who were not charged with cheating but who later admitted it had used this method to procure information.

One of the most staggering figures brought to light by the experiment was that relative to the incidence of cheating in the university where the study was conducted. Of the one hundred seventy-three subjects in the experimental group, ninety-six, or 56 per cent, were guilty of cheating on one or more university examinations. Only seventy-seven of the subjects, or 44 per cent of the group, were not charged with dishonesty. These are staggering figures, but they seem to be verified by other results of the study.

*The personality of the cribber.* The Bernreuter Personality Inventory was used as the instrument to secure a measure of the personality traits of the reagents. The blanks were passed to the students and they were allowed ample time to reply to the many items which comprise the test. This test gives a measure of the following personality traits: neurotic tendency, dominance-submission, introversion-extroversion, and self-sufficiency. To the extent that this tool measures these factors, it is believed that accurate results were obtained. No effort was made to create a situation for measuring ethical conduct in connection with this portion of the experiment.

TABLE II

COMPARISON OF CHEATERS AND NON-CHEATERS ON PERSONALITY TRAITS

| Factor | Cheaters | Non-Cheaters | Observed Not Admitted |
|---|---|---|---|
| B1–N score | —30.34 | —41.79 | —39.3 |
| B2–S score | — .01 | +16.26 | + 1.27 |
| B3–I score | —15.27 | —21.27 | —22.6 |
| B4–D score | +35.67 | +32.75 | +29.6 |

The cheaters were decidedly more neurotic than the non-cheaters, the differences in their scores being 11.45 ± 5.2 points. Although this finding is not extremely reliable, it is large enough to be considered significant. Furthermore, this is exactly what would be expected of those who crib

TABLE III

COMPARISON OF CHEATERS AND NON-CHEATERS ON VARIOUS FACTORS

| Factor | Cheaters | Non-Cheaters | Observed Not Admitted |
|---|---|---|---|
| Number of cases | 96 | 77 | 21 |
| Ages | 19.59 ± .31 | 21.05 ± .38 | 20.00 ± .49 |
| University grades | 2.94 ± .16 | 3.53 ± .10 | 2.05 ± .14 |
| Course grades | 3.14 ± .06 | 3.53 ± .10 | 3.14 ± .17 |
| Otis test scores | 42.27 ± .56 | 45.13 ± .72 | 37.48 ± .08 |
| Percentage of overstatement | 42.12 ± 3.8 | 30.82 ± 3.5 | 53.16 ± 2.6 |

on examinations, for, by definition, they are the ones who lean on others and who are somewhat deficient in emotional stability. In the matter of self-sufficiency, as measured by the Bernreuter Inventory, the exact reverse of the above results was found. Those who had procured illegitimate aid on examinations were far less self-sufficient than those who had not done so. The difference in this case was $16.25 \pm 4.1$ points, and was statistically reliable. When the other personality traits were considered, the differences were not found to be large, but the cribbers were found to be more *introverted* and more *dominant* than the non-cribbers. In all of these results, the mere size of the figures is not the only important factor; the abnormality of the groupings is worthy of further comment. Ordinarily, those who are neurotic and introverted are not dominant, and those who are dominant are usually self-sufficient. The fact that normal combinations were not found in the case of those who were guilty of unethical conduct may be offered as a partial explanation for their actions.

*The overstatement of the cribber.* The information overstatement test was constructed according to the pattern of a multiple-choice examination. On the first form of the test, the student was asked merely to indicate with respect to each item the degree of certainty he felt in his ability to answer it correctly. A three-point rating scheme was used, two points being given a question by the subject if he was certain that he knew the answer; one point, if he believed that he knew the answer but was not certain; and a zero score if he did not know the answer. The following are samples of the questions on this section of the test:

Do you know in which one of the four following years the Civil War began?
1848 1860 1861 1865 .....
Do you know which one of the four following metals a magnet attracts?
brass iron copper aluminum .....

The second part of the test consisted of questions of the multiple-choice type on which the student was expected to show whether or not he actually knew what in the first part of the test he claimed to know. The following are sample questions of this portion of the examination:

The Civil War began in
1848 1860 1861 1865
A magnet attracts
brass iron copper aluminum

This presentation of the questions in exactly the same form in the two parts of the examination represented an improvement over Raubenheimer's[2] technique, for with this procedure there could be no doubt in the subject's mind regarding what information was referred to by the experimenter. Moreover, in order to determine roughly to what extent a student might be unaware of his ignorance and misinformation, a control group was employed. The students in this group, in addition to being given the usual instructions, were informed that they would be called upon to demonstrate the information which they claimed to possess. Since, however, it was obviously impossible to be certain in any one case whether or not an individual's assertations regarding the possession

[2] A. S. Raubenheimer, "An experimental study of some behavior traits in the potentially delinquent boy," Psy. Monog., 34, No. 6, 1925, pp. 1-107.

of information he really did not have were intentional misrepresentations, no individual accusations were made. The data have been massed and reliance placed on such trends as comparisons between the groups of cribbers and non-cribbers revealed. The "percentage of overstatement" scores, used in Table III, were obtained by subtracting the average score made by the group on the performance test from the average score of the certainty estimates and dividing the difference by the performance-test average.

The scores for percentage of overstatement for the cribbers and non-cribbers were markedly different — 42.12 and 30.82, the difference being $11.30 \pm 5.1$. This is a fairly reliable difference, and it indicates that those who were guilty of cheating were prone to overestimate the extent of their knowledge. There is a possibility that these students actually believed that they possessed certain information, and, when they discovered their deficiency, they tried to cover their short-comings in any possible way. Both of these percentages are somewhat smaller than those Raubenheimer[3] found when he gave his test to elementary school children of superior moral quality, but the changed technique, which offered fewer opportunities for misinterpretation than his did, probably explains this difference. Furthermore, other factors, such as age, experience, and knowledge may have influenced the results of the two experiments in such a manner as to prevent fair comparisons from being made. In the special group, those who cheated but did not admit it, the percentage of overstatement was very large — $53.16 \pm 2.6$. This is nearly twice as great as that for the non-cheaters, and three times as great as that for the control group. Apparently these two forms of misconduct increase in a more or less constant ratio, and will be found together frequently.

*The self-scoring of papers.* The Otis Self-Administering Test of Mental Ability, Advanced Form A, with a twenty-minute time limit, was used not only to furnish an additional situation for testing deceit but to supply intelligence-test ratings. Following the procedure used by Yepsen,[4] the blanks were collected at the end of the allotted period, and an accurate record was made of each student's responses before the paper was returned to him for grading—a technique which made possible the detection of any subsequent alteration of a paper by a subject. In order to shed some light on the rôle of accidental errors in the marking and scoring of these tests, the control group was again employed. The students in this group were told, before the self-grading of the papers was begun, that an accurate copy of their papers had been made. Since, however, an earnest appeal for accuracy had been made to the experimental group, it is questionable whether the motivation for accuracy in scoring was any greater in the control group than in the experimental group. In any case, actual changes made by a student while scoring his paper would constitute unethical behavior.

Cheating on the Otis Test occurred among 45 per cent of those subjects who had an opportunity to score their papers. Twelve students were absent

3 A. S. Raubenheimer, op. cit.
4 L. N. Yepsen, "The reliability of self-scored measures," School and Society, 26: 657-660, 1927.

when this scoring took place, and, of this number, ten had been guilty of cheating on course examinations, whereas only two had not been so charged. Although somewhat less than one-half of the students were guilty of changing their Otis Test scores, this number was not divided equally between cribbers and non-cribbers. Of seventy-two who did modify their papers, forty-four, or 61 per cent, belonged to the group of cribbers, while only twenty-eight, or 31 per cent, of those who did not change their work had been guilty of cheating on course examinations. It should be remembered in connection with the former group that ten of the twelve students who were absent when the scoring was done had cheated on course examinations, and that this figure might easily have been increased materially had all students been present. It was found, in another portion of the experimental work, that those who were frequently absent from class were often found to be guilty of cheating when examinations were given.

There were reliable differences between the Otis test scores of the cribbers and the non-cribbers, the former having a mean score of 42.27, the latter a mean score of 45.13, the difference being $2.86 \pm .89$ points. This shows clearly that, on the average, those who cheated were below the mental level of those who did not cheat. Furthermore, the difference in Otis test scores for those who did and those who did not alter their papers was $6.70 \pm .84$ points. Here, again, it will be noted that those who were guilty were significantly less able mentally than those who were not guilty of unethical actions. There is reason to believe that a portion of the personality of the individual, both his attitudes and his actions, is the result of his native endowment.

In all cases except those of the regular examinations, where the spies were used, assurance was given the groups that their performance of the task at hand would have no influence upon any university marks or standings. In all instances the students seemed ignorant of the real purpose of the experimental technique, and they reacted spontaneously.

In the matter of scholastic achievement, those who were found to have cheated were well below the level of those who were not guilty. Reference to Table III will disclose the fact that the difference in the mean university grades for the two groups was $.59 \pm .11$ grade points. Furthermore, the mean course grade difference (this was the grade earned by the student in the class where he was observed cheating) was $.39 \pm .10$ grade points. It should be remembered, moreover, that a portion of this latter grade was earned by dishonest means. Both of these differences are very reliable, and both are decidedly against the cheater. When it is recalled that the cheaters were, as a group, below the mental level of the non-cheaters, it is only reasonable to assume that a portion of the pressure to which these students were subjected was the result of deficient native endowment.

*The non-admitted cribbers.* One of the most interesting comparisons made was that of the subjects who were observed cheating and who would not admit it with the rest of the group. Most of the subjects were quite frank in their admissions of dishonesty, but

there were twenty-one who denied being guilty, even when they had been observed in their misdeeds. In every comparison, these subjects appeared at a disadvantage when compared with the honest students, or even with the dishonest students who admitted their guilt.

The percentage of overstatement for this group was 53.16, or 22.34 ± 4.4 per cent more than the honest group. This is a very striking and reliable difference. Moreover, the mean intelligence test score for this group was 7.65 ± .74 points below that for the non-cribbers. This is a very great difference, being more than one and one-half standard deviations below the mean for the group as a whole. When the mean university grades were compared, it was found that this distinctly dishonest group had a grade-point average of only 2.05, as compared with a grade-point average for the non-cribbers of 3.53. This difference represents the difference between grades of "D" and "B–" on the normal grade curve. That this difference was of great importance cannot be doubted. The comparison between the mean course grades for the two groups did not reveal such striking differences, but it must be remembered that at least a portion of the grade in this course was earned dishonestly.

The difference was only .39 ± .10, but even this was extremely reliable. The tendency for this group to appear at a disadvantage in every comparison was very striking.

*Summary.* (1) Dishonest conduct, both minor and gross, occurred frequently among the students at a large state university. (2) The combinations of personality traits branded the cheater as a more or less distinctive entity, since the normal groupings of scores, according to the Bernreuter Inventory, were not found among this group. (3) All degrees of dishonesty appeared among the individuals studied — a fact which suggests that the categorizing of people into honest and dishonest is not justified. (4) The incidence of cheating in the classes under observation was a function of the pressure exerted by relevant factors and the ease with which misrepresentation could be accomplished.

# DANGERS AND VALUES IN TEACHER-MADE TESTS

D. Welty Lefever

ASSOCIATE PROFESSOR OF EDUCATION
UNIVERSITY OF SOUTHERN CALIFORNIA

*There is far more to test making than having them objective. Does the test diagnose teaching and learning difficulties? Does it enforce correct teaching practices? Does it differentiate between those who know the subject and those who are clever in guessing? Professor Lefever points out the difficulties, but he also shows how to avoid them.*

THE construction and use of objective tests is growing out of its adolescence and is coming of age. The novelty has worn off. Objective tests are no longer feared as an invention of the devil with which to mechanize the pupil's thinking and to thwart originality. They are no longer considered a panacea for all the ills of ineffective instruction. The testing movement is rapidly leaving the fad stage when teachers and administrators gave tests in order to be considered progressive and when little came of the testing except files of record cards, and a few neat looking charts in the superintendent's annual report. The present period is characterized by a sane and thoughtful recognition of serious limitations in the use of tests accompanied by an appreciation of important values.

So many problems are encountered in constructing homemade objective tests, and many of the difficulties involved are so easily unnoticed that a periodic review of the main problems and difficulties is valuable. The present article will attempt to bring to attention a few of the difficulties and dangers to be kept in mind in constructing teacher-made tests.

The number of items in a test which actually contribute to the capacity of the test to differentiate the members of a class into several levels of ability may be surprisingly small. Just because a test is apparently composed of one hundred questions or exercises is not proof that such a number of items enters into the measurement of pupil ability. If sixty items are so simple or well known that the group is practically unanimous in answering them correctly, the test in reality consists of only forty items. The writer was told of a course in which only one test was given. This test consisted of twenty questions of which eighteen were answered correctly by almost every one in the class. The marks for the course were based on two questions, not twenty.

Undoubtedly, serving as a basis for marks is not the most important use for a test, but obviously as long as marks are given, they should be as accurate as possible. Marks that are fair and just can result only when the class has been accurately divided into several levels of achievement or ability. Reliable differentiation is unlikely to occur unless there is a fairly large number of separate items assisting in the measurement of the range of ability in the group. If a considerable number of these items are too easy or too difficult, the measuring value of the test will be seriously impaired. A wise practice in this connection is to

keep tab on the number of pupils answering each item correctly. In this manner the difficulty of each question can be discovered.

It is important to balance a test with respect to the difficulty of the items composing the test. This means that if the questions or exercises are placed on a scale of difficulty as determined by the number of correct responses there will be only a few very difficult questions and only a few extremely easy questions. Most of the items will be of average difficulty. A test containing too many easy or too many difficult questions is likely to result in a badly distorted or skewed distribution of scores. If one test is properly balanced in difficulty for a highly selected graduate class in the university and if another test is balanced with respect to a high school freshmen group of the same size, the scores from the graduate group will approximate a normal curve as closely as will the scores made by the high school freshmen. A symmetrical distribution of scores means that the marks based upon such a distribution are more likely to be fair and true than if a badly skewed distribution were employed.

The major argument offered in favor of the new type test is that the scoring is more objective than when the traditional examination is used. A new type test may, however, appear to be more objective than it really is. This is a difficulty frequently overlooked. For example, a true-false test after the scoring key is once determined may be considered perfectly objective in the sense that a hundred scorers using the key would give the same test paper identically the same score. There may, however, be considerable disagreement among experts as to the correctness of the key. On going back into the construction of the test there may easily be argument over the wording of the items. Questions may arise regarding emphasis, choice of modifiers, the presence of specific determiners and the like which may cast considerable doubt on the final score. Still more disagreement could easily be found among the experts concerning the choice of items and the general make-up of the test.

The lack of agreement on the part of experts over the construction of a test would undoubtedly be much greater in certain fields than in others. Mathematics and science tests, in general, would probably result in less difference of opinion while in fields like sociology and education it is difficult to find many questions not purely factual or obvious in nature concerning which there would be complete agreement. In new fields of study like education and sociology, much of the objectivity of the objective-type tests is really a sort of "frozen subjectivity." The real test of objectivity is substantial agreement among those qualified to judge and not the mechanical application of a scoring key.

A valid objective test is one carefully constructed so that the emphasis given the items in the test will reflect accurately the emphasis received by the corresponding topics in the course. Otherwise the test is not measuring achievement in the course. Debatable questions will be eliminated at least in so far as they may directly affect the score. Ambiguous wording will be avoided. Instead, direct and simple statements will be used. No catch or

trick questions will ever be inserted to trip the unwary because that will make the test a measure of a rather shallow sort of cleverness rather than of achievement in the subject field.

One of the most difficult problems in test construction is that of steering the test between the Scylla of obvious questions on well-drilled subjects on the one hand, and the Charybdis of unimportant and equivocal questions on the other. If the important facts and principles of the subject have been given proper emphasis during the course most of the class should be able to answer correctly questions concerning them. This will result in few questions effectively measuring the range of ability in the group. On the other hand if the teacher searches in out-of-the-way corners of the textbook for obscure and unimportant items of knowledge, the student will be rewarded who memorizes his work in parrot-like fashion without distinguishing between the wheat and the chaff. Quite obviously a solution to the difficulty is the construction of questions requiring the student to exhibit in practical application a mastery of the principles concerned. These questions, in general, are more difficult to invent and may not always lend themselves readily to objective form. The writer believes that some considerable thought and practice given to the construction of this type of question will result in more effective examinations.

One rather valid objection to the more extensive sampling characteristic of the objective-type test is the emphasis usually placed on memory and purely factual information. Test items requiring of the student a simple recognition of surface facts are much easier to construct than those of any other type. For example, if completion questions are to be employed the temptation is to take key sentences from the text and to omit a number of words from each sentence. Unless one is very careful in the process words will be omitted in such positions in the sentence as to require the student to have memorized the wording of the text in order to answer the item correctly. Completion questions should demand of the student only a definite knowledge of the subject matter and neither the ability to recall verbatim portions of the text nor unusual cleverness at cross-word puzzles.

True-false tests are frequently guilty of skimming the surface of a subject. If a true-false question attempts to deal with more fundamental problems in many subject fields it is very likely to encounter controversy. An exercise in judgment for the student will undoubtedly have real value but it is questionable whether he should be penalized for answering a debatable question in a manner opposite to the point of view of the instructor.

The way in which tests are given and the character of the emphasis placed upon them by the teacher will determine to a large extent the study procedure likely to be followed by the student. Tests emphasizing the memorization of large quantities of subject-matter will tend to produce the parrot-type student who constitutes a large portion of our school population at the present time. Tests presenting problems requiring independent thought and reasoning based on the important principles of the subject will reward the thoughtful student and should stimulate genuine scholarship.

One problem in the mechanics of test organization and administration with which the teacher will have to deal is that of controlling the time element in the test. For most purposes the test should undoubtedly be of the power type, which means that a generous time allowance will be given the student. If the time set is too short the student will be forced to hurry over questions that should require careful thought and judgment. In this connection, the writer believes that the practice of dictating true-false questions is likely to be at fault. Unless the question is very simple and is presented in an unusually direct manner and unless it concerns itself merely with matters of information, there is danger that the student will not be able to give it sufficient attention to reply in a manner adequately representing his mastery of the subject. Students have frequently complained, and perhaps with justice, that so much of their attention was required to understand the wording of the question, that they did not have the opportunity to present their best thought on the question before being compelled to attempt the next item.

Teacher-made tests need not consist entirely of the so-called objective form for there is still value to be gained from the use of carefully prepared examinations of the subjective type. Some of the difficulties to be overcome in the organization of a well-planned essay-type examination are: (1) making the test questions clear and definite, so that the student knows whether to write a word, a sentence, or a short essay in response to the question, (2) employing questions which will compel the student to do independent thinking, (3) constructing the test so that it will constitute a reasonably extensive sample of the subject matter, (4) constructing a schedule of points to be covered in a complete answer to each question, so that the scoring can be much more definite and more reliable. The brief answer form of test has many of the advantages of both the old and new style examination. Its tendency is to require complete recall and not merely recognition as is the case with so many of the objective test forms. It is flexible and may be used to probe somewhat more deeply than surface information.

In spite of the difficulties and problems likely to be encountered in the construction of teacher-made tests, there are distinct values to be derived therefrom, as compared with the exclusive use of ready-made, standardized tests. The standardized examination seldom fits the course of study as organized in any particular school, and almost never emphasizes topics in the same manner as the instructor. If the use of standardized tests throughout the school is organized by the administration for the purpose of rating teachers, a great deal of harm is likely to result. Even where there is no formal attempt at rating, teachers may come to feel that their effectiveness is being judged on the basis of standardized tests which imperfectly match the work done in the classroom. The result is likely to be a tendency to teach to the test. If the results of the standardized test, as applied to the whole group rather than to the individual student, are taken too seriously, the outcome may be a tendency to stand-

ardize instruction and to prevent that constant adaptation of methods and materials which is the hall-mark of good teaching. Standardized tests, when employed as a basis for diagnosing pupil difficulties and when used by the teacher as a personal check on the effectiveness of instruction, become without question one of the most useful tools in teaching. The teacher-made test is probably of still greater value because of its flexibility, its closer approximation to the work of ~~of~~ the class, and because of its vital and personal relationship to the thinking of the teacher and the student.

# HOW TO STUDY FOR OBJECTIVE TESTS

C. C. CRAWFORD
PROFESSOR OF EDUCATION
UNIVERSITY OF SOUTHERN CALIFORNIA

*Cramming for subjective examinations means gathering together the tests used by an instructor with former classes and memorizing the answers to all the questions. Can one study effectively for the new type form of tests? There are many students who would like to know the answer.*

STUDENTS often ask, and usually in vain, for practical suggestions on how to prepare for true-false, completion, or other objective tests. They say they have learned how to study for the old-style essay test, but that the new tests present a new problem in study technique for which they are unprepared.

As a result of many such requests for help or advice from students, the writer has undertaken to assemble a few tentative suggestions. Thirty graduate students in a course on "Methods in Directed Study" were asked to submit any ideas they had acquired on this subject, either through personal experience or reading, and the best of their contributions are summarized in this article.

First will be presented general principles and policies that apply to studying for objective tests of whatever type, after which will follow more specific suggestions for the major kinds of objective tests.

## I. General Principles and Policies

The objective type of test, as a type, differs from the essay type in its demands upon the student, and the following suggestions may help the student to meet its peculiar requirements:

1. *Take the objective test seriously.* There is a common notion that it is easy, that guessing will yield a fair number of correct answers, and that if one question is missed, it does not matter because there are many others. The fact is that the test is as easy for other students as for oneself, and that a difference of one or two right answers may change the mark materially.

2. *Study by extensive methods instead of or in addition to intensive methods.* The objective test samples knowledge in a multitude of places instead of delving deeply into only a few topics. Furthermore, it demands somewhat more of the recognition

rather than of the recall response. This fact suggests the wisdom of being acquainted with a large number of aspects of the subject or course, even if lightly and superficially, rather than devoting too much time to a few topics. The soundness of this doctrine of extensive sampling is somewhat in doubt, we must admit, and deserves further investigation.

3. *Review for the test by skimming over the high points.* The reason for this policy is that instructors often make up their tests by turning the pages and hitting the "mountain peaks" of the subject; hence the student's need for something of the same process in matching his mind with the instructor's.

4. *Watch for key words, dates, proper names, and other highly concrete, factual, or objective items.* The demand for objectivity of scoring causes a teacher to gravitate naturally toward this type of questions, even though he realizes that he should test thought, appreciation, or other intangible products.

5. *Study the individual instructor's peculiarities and testing habits.* Some test makers stress facts, others opinions. Some select major points, others include more minor details. Experienced test makers have fewer trick questions or misleading items. Experience with an instructor may tell whether to answer a question as it is written or whether to look at it with suspicion to see where the catch or joker lies.

6. *Make up imaginary test items beforehand.* The student's selection and statement of items may not match the instructor's very closely; yet the effort to do so will bring about the formation of many of the same bonds as will be required for the instructor's test. The plan of students giving each other objective tests is commonly reported, and those who have tried it seem convinced of its effectiveness.

7. *Study any available test questions on the subject.* Tests used for previous classes, tests given by other instructors, or tests printed in books, study guides, etc., upon the subject, all give help and guidance. Two tests may not have a single item exactly identical; and yet each will probably stress somewhat the same fundamental issues.

8. *Read about the theory of objective test construction.* The instructor is the fisherman and the student the fish, as it were. From the standpoint of the fish it would be advantageous for each fish to read a good treatise on the fisherman's technique in order to avoid being caught. It would at least be interesting if someone would compare objective test scores of students who have had such instruction with those of students who have not.

## II. Specific Procedures and Policies for True-false Tests

The true-false test statement, by its very nature, has an individuality of its own. The following suggestions were compiled as contributions to study technique for this type of test.

1. *Hunt for debatable issues.* The test maker is likely to choose issues which are fairly close or debatable, rather than those which are obviously true or obviously false. Popular superstitions, common misconceptions, and points upon which the man in the street differs from the scientist or scholar are especially likely to be used for true-false items.

2. *Get the viewpoint of the author or instructor.* Since the test is to be

scored in the light of the source from which it is drawn, the student must know which side of a debatable issue the author or instructor takes. He may disagree with either or both, but can at least answer the test according to the book or teacher. A true-false test paper is a poor place to insist upon one's own views when they are in conflict with those of the authorities on the subject.

3. *Check the book or notes very carefully for positive, emphatic, epigrammatic, or startling statements.* The instructor is likely to select test items from this type of points instead of resorting to more prosaic or commonplace issues, because they catch his attention when he turns the pages hunting for likely questions, and also because they are already practically ready-made true-false items.

4. *Note especially the exceptions to general rules and principles.* The test maker often wishes to test knowledge of a fundamental point which if stated as such would perhaps be too easy; hence in order to make it into a valid true-false item he "slips up on it from behind" by introducing one of the exceptions.

5. *Get the fundamental viewpoint or philosophy of the subject.* The specific statement is often apparently trivial or unimportant in itself but is included because it illustrates or relates to something much more fundamental. Students sometimes get the idea that true-false tests do not amount to much, and hence do not study effectively for them, because of this failure to recognize fundamentals behind examples or test cases.

6. *Note any specific clues in the wording of the statements.* Such words as always, never, everyone, no one, etc., furnish valuable clues suggesting false statements, while words like often, usually, or generally, soften a harsh statement down and make it much more likely to be true. An expert on test construction, G. M. Ruch, has found that one can materially improve his score on a typical true-false test on an unfamiliar subject by noting such special words, which he calls "specific determiners."

7. *Avoid extremes of doubt and suspicion.* The fact that some statements are false does not mean that all are tricky, misleading, and intentionally deceptive. Some instructors, it is true, resort to trickery and deception because they have not yet become expert in test construction, but most tests are plain, straight, square, honest, and above board. The student who suspects an ambush at every turn is likely to do worse than the one who reacts to each statement as if it really meant what it appeared to mean.

### III. Specific Procedures and Policies for Completion Tests

The completion test does not have a very distinct study technique, so far as could be determined in this investigation, but such tentative suggestions as could be collected are presented in the next few paragraphs.

1. *Study to recall rather than merely to recognize points.* The true-false test is a mere recognition test, but a completion test requires a more thorough mastery. It is well, therefore, to test oneself for ability to name the points that are likely to be asked, if the test is to be of the completion type.

2. *Memorize any available lists or outlines of points.* A common tendency of test makers is to formulate questions that call for enumeration of

the subdivisions of an outline, or that call for supplying the general category in which a specific subpoint falls. Any definite list or outline should therefore receive extra attention.

3. *Specialize on new words or technical terms.* A word that is peculiar to a restricted subject or topic makes an unusually good basis for a completion test, since its meaning can be stated as a test item and the word itself left blank for the student to fill in. A student who has a good record in completion tests reports that he underlines all new words as he reads, and later gives them a very special review.

4. *Concentrate on definitions, laws, dates, and other highly definite items.* The difficulty of scoring completion tests with a high degree of objectivity leads the test maker to select a large percentage of his items from the "dry bones" of knowledge and memory rather than from the less tangible products requiring thought and judgement. Names and places may not be as important as principles and concepts, but they are more easily scored.

## IV. Specific Procedures and Policies for Matching Tests and Best Answer Tests

The sources from which suggestions were drawn in this study seemed not to be very fertile with respect to study techniques for the matching and best answer tests. Mention was made of the importance of understanding important terms, of hunting for any natural pairs of points, such as cause and effect or men and events, and of getting a real mastery of important concepts and principles so that their relationships to other items could be readily recognized. Since matching tests very often measure a recognition of the part-whole type of relationship, it is suggested that considerable effort be made to get a rich fund of detail or associative content to go with each major point, principle, or concept. The best answer test is psychologically about the same as the matching test, in fact a simplified form of matching, since answering requires the selection of one item from several and making it agree with the original statement.

## V. Some Final Considerations

The meagerness of the offering presented here suggests either that there is little that students can do to improve their scores in objective tests through the study techniques employed, or else that if there are many really helpful methods of work, few students have discovered them. Perhaps the difficulty here encountered in our efforts to formulate study helps for students in taking objective tests points to a needed extension of the efforts of the test and measurements experts.

Perhaps, on the other hand, the fact that students know so little about how to "beat the game" in taking objective tests is cause for congratulation. Perhaps it is just as well that we leave them in ignorance of the fine tricks of the test taker's profession, and let their papers reveal knowledge of the subject matter of the courses taken rather than knowledge of the clever tricks of getting a better mark than they deserve. At least, some professors see the situation this way, and perhaps they are right. At least, if students are to be coached in methods of taking tests, the coaching should be made available for all instead of for just a few who get it through "bootleg" channels.

# SOME RESULTS OF THE TESTING MOVEMENT

J. HAROLD WILLIAMS

PROFESSOR OF EDUCATION
UNIVERSITY OF CALIFORNIA AT LOS ANGELES

*Too long has education been dominated by theories which have not been valid. Measurement has shown how to improve teaching practices and methods of organization. It has been a determining factor in curriculum revision and adjustment of the program to meet individual differences. It is probably not an over-statement to assert that measurement is as important in education as accounting is in business.*

AT a meeting of the National Education Association in 1897 Dr. J. M. Rice suggested that spelling ability in children might properly be measured by giving them a list of fifty selected words and counting the number correctly spelled. It had been his chief purpose to discuss the relative merits of different methods of teaching; but his suggestion of a plan of objective measurement, with more or less standardized procedure, was a new idea, and one which has been of profound influence in education.

There was ample precedent, at the time of Dr. Rice's experiment, for applying objective methods in the systematic study of facts. Progress in other fields of human endeavor, notably the physical sciences, had been made in almost exact proportion to the extent to which measurement methods had been made available. The making of measurements still occupies the time of many workers in every science. Evaluation is a part of complete observation.

There was also precedent for the protests which arose from some educators against any attempt to measure educational results, or to express such results numerically. The early scientists, seeking more exact methods of observing and evaluating natural phenomena, were obliged to meet opposition of the same sort. The fear was expressed that measurement was too materialistic; that it was an undignified procedure, unbecoming persons who were capable of reflective thinking; and that the examination of factual data might divert the mind from more important philosophical considerations. But the subsequent development of astronomy, physics, and chemistry seems to have justified the objective approach. Today the leading concepts of world structure are based upon evidence obtained from measurements; and in education such procedures have been among the most effective means of improving instruction and school organization.

Educational measurements compare favorably with physical measurements in their applicability and relative accuracy. Thorndike has pointed out that certain physical standards, such as absolute zero, are of very recent determination; and that courts of law make important decisions involving the use of physical units which are no better established than are many of those in regular use in connection with the interpretation of standard educational tests.

Improved teaching methods have followed closely upon the testing movement. Although successful teach-

ers have always known that learning depends to a certain extent upon intellectual development, it is only since standardized tests have become available that the relationship could be observed or utilized in a practical way. Children who under older practices might have been given up as being unable to learn, but who are shown by intelligence tests to have mental capacity beyond their present achievement, may safely be encouraged and expected to do better. On the other hand, children of limited mental capacity, who under traditional systems might be prodded to the extent of discouragement, may now be assigned programs within their mental grasp, and credited with achievement which is normal for them. Without the use of measurements extending beyond the scope of ordinary class-room observations such adjustments are not easily made. Probably some of our best teaching is accomplished in this way with segregated groups of children from the extremes of the distribution curve; the subnormal and the gifted. The normal child also has better opportunities in the school in which test results are regularly applied.

The principle that school achievement can be measured in definite units, and in the light of norms, has given instruction an impetus which was much needed. School surveys repeatedly pointed out that schools and school systems differed from each other greatly in their accomplishments, even with similar time allotments; and that these differences were often due to the failure of some schools to make use of available standards of achievement. The deviations were not all on the negative side. Some schools were actually achieving more than was necessary or desirable in certain subject matter fields, while neglecting important activities. In these cases the systematic use of tests has been recommended as an economy measure.

Not only standard tests, but new-type tests of all kinds came into use when it became evident that improved examination methods tend to produce better instruction. Thinking teachers saw in this method an opportunity to eliminate much of the guess-work that had characterized their previous efforts. School systems not only made regular testing a part of the school procedure, but also embarked upon programs of test construction, with the establishment of local norms to accord with their own curricular needs. In many cities this work has been carried on with great care and accuracy. It has become concrete evidence of an earnest desire on the part of school officials to conduct their schools efficiently.

In the promotion of pupils the consideration of test results has become a common practice. This is particularly true where promotion means transfer to another school, as from elementary to junior high school, or from junior to senior high school. Even when traditional grading systems are retained, it has been deemed wise to distribute marks in such a manner as to avoid great deviations from test findings. This has been especially desirable where school marks are interpreted as evidence of ability to pursue a course of higher education. Secondary schools now feel this responsibility more keenly than ever. Most of the recent improvements in marking systems are based upon evi-

dence secured from test results. These proposals usually include the use of standardized procedures, and the elimination of the subjective element in judging abilities and accomplishments. It seems probable that the whole problem of issuing grades and marks will come in for some serious discussion during the next few years, and measurements will be found necessary to such considerations.

Test results are especially useful in the study of plans for curriculum revision. Although occasionally the fear is expressed that formal testing may in some way bring about too rigid adherence to traditional subject-matter, it appears that in reality the opposite effect is being produced. Test construction leads to a better evaluation of teaching aims and objectives. Curriculum revision has been most active during the testing era, and has been greatly stimulated by measurement investigations. These two types of research have often gone hand in hand as their interdependence has been recognized. Some of the most significant changes in curricula have been brought about in school systems in which measurement facilities were available for checking the results of curriculum experiments. It does not accord with the spirit of this scientific age for anyone to claim that progress has been made in a given direction without producing some objective evidence of the asserted change; and most curriculum makers have shown a disposition to avoid such errors. Experiments with the activity program, for example, have included the use of standard tests in the traditional subject-matter fields, with the subsequent finding that the newer plan of organization contributes fully as much to achievement in these subjects as any other plan of instruction. Other experiments have shown that some methods of teaching are about twice as effective as others, as judged by relative achievement. Without test results such claims would rest upon very uncertain ground. For purposes of this sort standard tests furnish one of the most reliable forms of evidence.

The more intensive study of the curriculum has made it necessary to state the objectives and methods of education, at least for practical purposes of school organization, in more definite form than has been the practice of former years. The test movement has contributed to this purpose a terminology which the schools have been quick to adopt. Such terms as mental age, I. Q., subject age, norm, prognosis, correlation, grade-placement, validity, reliability, T-score, occur in educational discussion without the necessity of special explanation. At the same time certain other words and expressions which are not new to education have taken on new and more definite meanings, in consequence of their use in testing programs. Among the latter are such words as achievement, development, guidance, distribution, survey, gifted, accelerated, retarded, maturity, homogeneous, capacity, tabulation.

Tests have contributed more than any other method of observation to the study of individual differences. To give every child an education suited to his capacity is an aim which could hardly be realized without the use of objective tests. Beginning with the segregation of retarded and subnormal children the differentiation

of instruction has reached the point of complete ability grouping in some schools. Although practical difficulties have been encountered in making such a plan work in schools which have long been organized on a traditional basis, the success obtained in some schools has aroused much interest and discussion. There is little tendency, in all of the opposition to ability grouping, to refute the contention that learning is most efficiently accomplished in a situation appropriately gauged to the ability of the learner. If homogeneity is not possible in all classes, schools have undertaken to determine where it can be obtained, and under what conditions it is most effective. Test findings play an important part in the study of the problem.

The measurement program of today's school is not a haphazard utilization of experimental tests. It is a carefully planned procedure, usually directed by a well-trained person, and will be found to contain a balanced diet of selected standardized tests, the validity and reliability of which have been established. It includes the use of a series of achievement tests, appropriately fitted to the curriculum and local needs, and suitable tests of intelligence, aptitude, interests, and emotional stability. The extent of the program depends partly upon the funds available, and these are being wisely spent; it also depends partly upon the attitude of the school administration toward the use of objective methods, and the encouragement given teachers for their practical application. Under such a program every child in the school system will have the benefit of measured results. Teachers will have something definite to supplement their own opinions in matters of promotion and grading; parents may have more confidence in the work of the schools, because of the elimination, in part, of the subjective elements which too often influence school reports. Administrative officers may feel more secure in the use of results obtained by test procedure, for there is no part of the school system which is based upon sounder investigation. The single fact that such results make possible the comparison of one school system with another justifies the use of standard tests.

The school of the future may be characterized by a child-centered, naturally-motivated, activity program; it may be even more rigid and mechanical in some respects than are the schools of today; it may be both of these. In any event it appears self-evident that there will be an increasing need for developmental facts, obtained by reliable methods of observation; there will be more need than ever for methods of measuring what is being accomplished. Enlightened parents will seek more information about what the schools are doing, and what changes are being brought about in their children. Taxpayers may demand from the schools an accounting of results, in terms of objective standards which will be intelligible to the layman. All of these require not only the application of what has been learned with the use of tests, but the making of more and better tests. It is inconceivable, in the light of recent progress, that education should be again dominated by theories which are not supported by objective evidence.

# THE EFFECTS OF REST PERIODS ON PROGRESS IN LEARNING IN TYPEWRITING

A. S. Raubenheimer and S. H. Davidson
UNIVERSITY OF SOUTHERN CALIFORNIA

*The authors find scientifically what most teachers have generally felt to be true, that interruptions in the learning process are detrimental. They also find, what many teachers who assign heavy amounts of home work need to have brought to their attention, that in typewriting at least, irregular, unsupervised work in the beginning stages of learning is definitely detrimental to progress.*

IF educational efficiency is to be attained, the effect of recesses upon the learning process along with other factors must be ascertained. Data bearing upon improvement or deterioration of the stimulus-response bonds after recesses are very limited in quantity. Those experiments which have been performed have dealt with a few selected pupils and have been conducted under the carefully controlled, yet artificial, conditions of the laboratory. More experimentation, therefore, is needed, and, in particular, carefully conducted experimentation in specific fields in the practical circumstance of the classroom.

*Statement of the problem.* The problem of this study may be stated as follows:

1. It aims to determine to what extent, if any, speed of typewriting is increased or decreased by rest periods, varying in length from two to fourteen days.

2. It aims to determine to what extent, if any, accuracy of typewriting is increased or decreased by rest periods, varying in length from two to fourteen days.

*Related investigations.* The bulk of the evidence pertinent to the deterioration or improvement of stimulus-response bonds following periods of disuse has been compiled by Thorndike in his *Educational Psychology,* Volume III. Thorndike asserts that, in general, periods of disuse are followed by a weakening of the stimulus-response bonds and that, other factors being equal, this weakening is proportional to the length of the period of disuse. Of the experimental studies summarized by Thorndike, two are pertinent to the present investigation:

1. Book reports a study in which Individual A was tested for ability in typewriting after long periods of disuse. "A" averaged 1508 strokes and 2.21 per cent error in ten 10-minute tests from January 7-16, 1906. After a long rest period he was retested from June 1-10, 1906, averaging 1433 strokes and 3.54 per cent error. After a still longer rest period, he was tested daily from June 1 until June 10, 1907 and averaged 1611 strokes and 3.15 per cent error.

2. Rejall used himself as a subject in studying the permanence of ability to typewrite. At the end of 30 hours of practice, he could type 25 words per minute with 4 errors per hundred words. After a rest period of 3½ years, he made the following scores on the first five days:

18.7 words per minute with 8 errors per hundred words.

18.9 words per minute with 7⅓ errors per hundred words.

21 words per minute with 6⅔ errors per hundred words.

22.1 words per minute with 5 errors per hundred words.

22.5 words per minute with 8⅔ errors per hundred words.

By continuing the practice he regained in five additional hours of practice nearly the same efficiency that he had obtained in 30 hours, three and one-half years before.

Other experiments summarized by Thorndike involved the effect of periods of disuse on such mental functions as (1) memorization of nonsense syllables (Ebbinghaus and Rasossawljewitsch), (2) ball tossing (Swift), (3) multiplication (White) and (4) addition (Kirby). As Thorndike himself states "the only facts that they display with any unanimity and brilliancy are the apparent complexity and variability of deterioration by disuse and our lack of knowledge about it."

*Procedure.*

1. The subjects used in this study were pupils in regular typewriting classes at the Huntington Beach Union High School during the school year 1928-29. Seventy-five students enrolled with the same teacher in three sections of first-year work were included.

2. All of the first-year students were given regular classroom instruction in accordance with the Weise-Coover System of teaching typewriting. The daily period was 55 minutes long.

3. The first-year students were given daily 15-minute speed tests, beginning October 15. Test material was selected from the Wiese-Coover text, "Kinesthetic Method of Touch Typewriting" and from test material prepared by various typewriter manufacturers.

4. Exact records of test results for each student were kept, the following items being recorded: number of test, date of test, gross number of words typed, number of errors, net number of words after deducting ten for each error, gross rate of strokes and the net number of words per minute. The records were kept on "Pupil's Daily Record Sheet" forms, provided by the H. M. Rowe Company, for use in schools using the Wiese-Coover method.

5. Tests given before and after the following rest periods were carefully studied for decreases or increases in speed and accuracy:

October 20-21—Two days
October 27-28—Two days
November 3-4—Two days
November 17-18—Two days
November 24-25—Two days
December 8-9—Two days
December 15-16—Two days
November 10-12—Three days
November 27—
December 2—Six days
December 19-January 1 — Fourteen days

6. Since it would be impossible to require that all pupils refrain from practice during rest periods unless they were told of the study and its purposes and since it was thought that such knowledge might cause an unnatural strain during tests given before and after rest periods, it was decided that all students should turn in practice slips showing the number of minutes spent in practice each day. These slips were required from all students, but it was made clear that outside practice was not a requirement. Such slips were initialed by the parents.

*Analysis of results.* Comparisons were made of average gross speeds, number of errors, per cents of error,

and net speeds before and after rest periods.

The average gross speed was 274.5 words per 15 minutes before the two-day rest periods and 275.3 words per 15 minutes after these periods; the average gain in gross speed being only .8 word. The increase of 178 words in gross speed obtained in the 33 learning periods between October 20 and December 17 represents an average gain of 5.4 gross words per learning period. Thus the average gain in gross speed over a two-day rest period is only one-sixth as great as the average gain for the normal learning period. After the three-day rest period there was observed an average loss of 18 words in gross speed; after the six-day rest period there was a 2-word average increase; the fourteen-day period was followed by a 10-word average decrease.

Following the two-day rest periods, testing revealed an average decrease of .5 errors. There was an average increase of 1.9 errors after the three-day rest period, of 4.4 errors following the six-day rest period, and of 9.7 errors following the fourteen-day rest period.

There is a decrease in the average per cent of error following the two-day rest period, from 3.3 to 3.1 per cent. After the three-day rest period it increases from 3.0 to 3.9 per cent, after the six-day rest period from 3.3 to 4.7 per cent, and after the fourteen-day rest period from 2.6 to 4.2 per cent.

A consideration of average net speeds revealed that there was a gain of 8.2 words following the two-day rest periods. This gain is considerably in excess of the average gain of 4.4 words in a normal learning period. Following the three-day rest period there was an average loss of 33 net words, following the six-day rest period an average loss of 42 net words, and following the fourteen-day rest period an average loss of 60 words.

Table I gives the average number of gross words, average number of errors, average per cent of error, and average number of net words before and after rest periods.

On studying the records of those pupils who engaged in irregular, unsupervised typing practice during the rest periods, that is, at home over week-ends, it was found that by far the majority showed losses both as to speed and accuracy.

*Conclusions.* The study seems to justify the following conclusions:

1. All rest periods which are at least two days in length tend to retard progress in the learning of typewriting as measured by gross speed; the retardation in gross speed increases with length of the rest period.

2. Following a two-day rest period there is a slight improvement in accuracy; following longer rest periods there is a loss in accuracy, the loss increasing with the length of the rest period.

3. The two-day rest period may be considered beneficial to the learning process since the increase in net speed indicates that the improvement in accuracy more than offsets the retardation in gross speed.

4. Rest periods of greater length are detrimental to learning since there is a retardation in both speed and accuracy.

5. Irregular and unsupervised typewriting by pupils in the beginning stages of learning is definitely detrimental to progress in learning.

# NEEDED DEVELOPMENTS IN MEASUREMENT FOR SECONDARY SCHOOLS

J. Murray Lee

DIRECTOR OF RESEARCH, BURBANK, CALIFORNIA

*The fifteen years that have passed since standardized objective testing was first introduced into the secondary school have shown mainly how much more needs to be done. It is still virgin territory and many a doctor's dissertation will be found here.*

WHAT are the next steps in the program of tests and measurement for secondary schools? Occasionally in any work it is necessary to take inventory and project plans for the future. Standardized testing in the high school began about 1916.[1] Since that time tests have been published in increasing numbers until now there are over 610 tests[2] commercially available for use in the secondary schools. According to the best estimate available there were over 6,000,000 standardized intelligence and achievement tests used in the secondary schools during 1931.[3] It is not the purpose of this article to evaluate this movement or to show trends[4] but to set forth needed developments as suggested by public school workers, namely, high school principals and research directors.

The material which is included has been gathered from two sources. The first source was the incidental comments appearing on a questionnaire to high school principals requesting information concerning testing practices in their schools. Replies were received from nearly 500 schools representing all forty-eight states. The second source was from a check list which was filled out by 106 research departments from all parts of the United States. Among other data the check list requested the changes which the directors would suggest in their own department and a statement of what they felt were the most needed developments in the field of testing for secondary schools. In analyzing these data there has been no attempt to show the relative importance of each problem discussed. The purpose has been to present the thinking of persons in the field who are dealing with the actual situation and in their own words whenever possible. It is felt that such a treatment is much more valuable than presenting the thinking of one person in this field.

## Opinions of Administrators

While there was no definite question on the check list to high school principals as to needed developments, many interesting comments appeared on the check lists. These comments when summarized seem to offer valuable material when considering needs in this field.

The most outstanding problem, as might be expected in this period of decreased budgets, was the lack of funds and also personnel. Some of the statements bring this out very clearly.

1 J. Murray Lee, "Testing Programs for Secondary Schools." An unpublished doctor's dissertation on file in the Teacher's College Library, 1932, Chapter II.
2 Ibid.
3 Ibid., Chapter I.
4 David Segal, "Trends in Tests," School Life, Vol. XVIII, Oct., 1932.

"Due to lack of funds, no tests will be given this year. Results could be used to a greater advantage if adequate administration and supervisory personnel were available."

"We have no program because of lack of funds. Principal does what he can with no Board of Education support."

"Our system makes no allowance for time or help in testing. Cost and extra time allowances have been our largest handicaps."

"We have no testing since the depression hit us."

Another statement reflecting the favorable situation when trained personnel is available follows.

"We are just commencing to use our test results in guidance, since we now have counselors who are able to use them."

Several of the schools, including the one just quoted, were just beginning to use tests or to desire them.

"The program in this school is just being organized and present methods are somewhat confused. A complete educational, physical, and personality study, as well as a social history, will eventually be available for all children in the school."

"Very little testing has been done in the high school. We are planning, hoping to do more."

One principal from a large city, having a research department, writes:

"No testing program has yet been established in the school. I am not satisfied with the situation."

Expressing an opposite point of view, another says:

"I wish we might have more real teaching and less effort spent in petty testing. We have junked most of our testing program and I think we are ahead because of it."

Other statements stress the use of teacher-made tests.

"The tests are planned here in the departments and changes are made each year to meet the needs of the curriculum. . . . I find that an imposed program of testing does not arouse the same support that is given one worked out by the teachers interested."

"We are doing quite a good deal of studying and developing teacher-made comprehensive tests."

"We have had teachers from each school in the city plan tests for the entire city. In this case there was stimulated considerable interest."

Another trend indicated was away from the use of intelligence tests. Some of these statements were most interesting.

"We use the achievement test exclusively. It correlates very closely with the so-called intelligence test and it has not its implications."

"Our use of tests is not what it should be, due to the fact that our Board of Education has ruled against the use of intelligence tests throughout the system. We find that percentile ranks on a reading examination give us a fair substitute for the intelligence test score."

"We are not very strong for intelligence tests. A good achievement test will show everything that an intelligence test is supposed to show."

And finally from one principal comes:—

"I find mental tests utterly unreliable as given. Unseasoned people try to determine infinite questions with finite means and fail, as they should. Am absolutely disgusted with the propaganda of psychic mountebanks. They are reverting back to the system of thought that the school men used in the middle ages."

These quotations should be considered as each representing only one opinion. They do, however, illustrate the extreme feeling at both ends of the scale on testing and indicate several needs of the secondary school. These needs are: 1. Money for the

purchase of tests. 2. Trained personnel to handle the testing programs. 3. Teacher-made tests which will measure the outcomes of instruction. 4. Principals with more knowledge of modern testing.

### Needed Improvements in Organization is Expressed by Research Directors

Most of the directors did not reply to the question concerning the needed changes in the organization of their department or made the statement that present conditions were satisfactory. From the 20 percent which suggested changes, several significant trends are apparent.

First, an improvement in the use of test results by the teachers.

Second, the need of more work in the actual construction of tests.

Third, an enlargement of the work on the secondary level.

Fourth, a program for the improvement of teachers in their ability to make good tests.

Fifth, greater cooperation and interest on the part of teachers and principals.

### Needed Developments in the Field of Secondary School Testing as Expressed by Research Directors

There were thirty-four different items mentioned as being the most needed developments in the field of testing in the secondary schools. The largest number mentioning any single item was ten. This offers an indication that there is not one single problem which is pressing but a rather large number. Frequency of mention is not to be taken as an index of importance for probably the replies were given very quickly and they had no list of problems to guide them. The list does contain a large number of problems which are in need of further consideration by research workers in general and does give an idea of the type of problems which research directors feel need solutions.

*Better Tests*—Better tests are needed. The suggestions for improvement which were given were, that tests should be more reliable, have 5 or 6 forms, be shorter and be more thoroughly standardized. Short tests which are reliable must cover a fairly limited body of subject matter, while more thorough standardization probably should be undertaken by the publishers. The expense of putting a test on the market does not lie in its publication but rather in its standardization. Until the publishers are willing, to a greater extent than they have been, to assume part of this burden, better standardization of tests is difficult. Needless to say, the responsibility goes back to the purchaser to insist upon adequate standardization.

*Training of teachers* — The comments stressed need for teachers to be better trained in the use of tests, in making valid tests, in a better understanding of what tests reveal, and in their values in different situations. The opportunity for this training has largely been through college work and a surprising per cent of the teachers report having had courses in measurement. There would seem to be an opportunity here for training in service which a goodly number of research departments and principals are neglecting.

*Constructing new type tests* — The emphasis was placed on the construction of tests to fit the local course of

study. The following quotation is typical:

"More tests made in system to fit subject-matter or courses of study benefits teachers as well as makes better tests."

*Prognosis or aptitude tests* — More prognosis or aptitude tests are needed for guidance in the selection of subjects and in the discovery of special aptitudes. The opinion of one director is:

"Prognosis tests for guidance purposes here are one of the greatest needs."

Such tests are available now for Algebra, Geometry, Latin, modern foreign languages, commercial subjects, art, music and mechanical ability.

*Outcomes other than factual*—This field offers large opportunity for work. The directors desire tests which will "measure attainment of educational aims other than information" and "other than mere knowledge," and those which "test thinking capacity as well as information." One member of a research department states:

"I feel that when such techniques (the measurement of other outcomes) can be developed, the whole movement for the improvement of measurement in secondary education will go forward much more rapidly than it has in the past."

The others say that they wish:

"A technique for testing objectives and outcomes other than content and factual knowledge," and "a new type of test that measures significant relations as well as facts."

The latter "is especially needed in literature and social studies."

It seems significant that the six directors quoted above, are in cities with population above 150,000 while four of them are in cities larger than 400,000. It is hoped that these directors in the more important positions will contribute to the solution of this problem.

The remaining developments which are given in Table I were mentioned only three times or less. The items which are lettered were given only once and they are direct quotations from the check lists.

This list of needed developments should be of value to research departments in considering where others feel problems exist, to teachers of measurement courses in trying to make their courses more functional in solving the problems of the field, and to research workers in selecting problems for which there is a felt need for solution.

TABLE I

A LIST OF NEEDED DEVELOPMENTS IN SECONDARY SCHOOL TESTING AS REPORTED BY RESEARCH DIRECTORS

| *Rank* | *Frequency of Mention* | ITEM |
|---|---|---|
| 1.5 | 10 | Better tests. |
| 1.5 | 10 | Training teachers in making and using tests. |
| 3. | 9 | Constructing new-type tests based on the local course of study. |
| 4. | 8 | More prognosis or aptitude tests. |
| 5. | 6 | Tests which will measure outcomes other than factual. |
| 8. | 3 | More tests. |
| 8. | 3 | More diagnostic tests. |
| 8. | 3 | Better tests for vocational guidance. |
| 8. | 3 | More reliable norms. |
| 8. | 3 | Tests based on desired outcomes of instruction. |
| 11.5 | 2 | Comprehensive batteries with comparable scores. |
| 11.5 | 2 | Diagnostic reading tests. |

# EDUCATION IN 1940*

FRANK C. TOUTON

VICE-PRESIDENT OF THE UNIVERSITY OF SOUTHERN CALIFORNIA

*To one who sees school buildings erected in 1933 from plans that were obsolete many years ago and who finds teachers being hired on a basis of relationship to some member of the School Committee rather than the ability to teach, the thought comes that Professor Touton's 1940 may need to be changed to a later date for some of the items he has included.*

## *Introduction*

A BACKWARD look over four decades in education reveals changes and progress which no conservative educator of 1890 would have predicted. From a review of certain phases of this development it is my purpose here to mention some features of the educational program which will appear, I think, at the end of the present decade.

The items considered are: growth in school enrollments and in the number of teachers employed in all levels of instruction, preparation of the teaching staff, purposeful education, school classrooms, methods of learning, use of moving pictures, student activity in the classroom, and equality of educational opportunity.

*Growth in School Enrollments.* Prior to 1890 enrollment figures were not gathered as now by the Office of Education of the Department of the Interior. Reports earlier than 1890 do not provide comparable data and are not here given. The figures given below are from Bulletin No. 20, 1931, of that Department. The predictions for 1940 in the several divisions are approximations, made with due regard to the trends of enrollment exhibited for the ten-year periods, 1890 to 1930 inclusive. Several factors are judged potent in determining the projected trend for the decade from 1930 to 1940, among which are the restrictions on immigration and the lowered birth rate, which will certainly operate to decrease the rate of increase in elementary school enrollments. The present general unemployment almost certainly will operate to raise the compulsory age limit of school attendance, in order that the places of children in industry may be filled by now unemployed adults. Then, too, the pressure of our economic situation will retard the usual rise in enrollments in the upper grades of the secondary schools and in the several levels of higher education, both public and private. The growth of professional school enrollments cannot be expected to continue at its present upward rate, for the professional services which their graduates are prepared to render must bear a constant ratio to the present reduced rate of growth in the population unless new services are discovered and trained for by professional schools.

Below are given in thousands the school enrollments by decades since 1890, with predictions for 1940.

* A lecture given by the writer on April 16, 1932, at the Western Pennsylvania Educational Conference, Pittsburgh; and on November 28, 1932, at the University College of the University of Southern California.

*(Enrollments and Predictions are Given in Thousands)*

| *School Level* | *1890* | *1900* | *1910* | *1920* | *1930* | *1940* |
|---|---|---|---|---|---|---|
| Kindergarten | 32 | 225 | 346 | 511 | 778 | *900* |
| Elementary and Kindergarten | 14,181 | 16,224 | 18,457 | 20,864 | 23,588 | *26,000* |
| Secondary | 358 | 696 | 1,111 | 2,495 | 4,800 | *7,000* |
| Normal Schools and Teachers Colleges | 35 | 70 | 89 | 135 | 162 | *175* |
| Colleges, Universities and Professional Schools | 158 | 168 | 267 | 462 | 924 | *1,250* |
| Special Schools | .. | .. | .. | .. | 500 | *700* |
| Junior Colleges (included above) | .. | .. | .. | .. | 75 | *125* |

The data on enrollments and predictions of the above table reveal the following trends:

(1) Since 1890 the kindergarten enrollments have increased from 32,000 to 778,000 in 1930 and will probably reach 900,000 in 1940.

(2) Since 1890 the kindergarten and elementary school enrollments have increased from 14,181,000 to 23,588,000 or approximately 64% and will probably reach 26,000,000 in 1940.

(3) Since 1890 the secondary school enrollments have increased from 358,000 to 4,800,000 and will probably reach 7,000,000 in 1940. During the past forty years the secondary school enrollments have practically doubled each decade, a growth which is not matched by that of any other social institution of like magnitude anywhere in the world.

(4) Since 1890 the normal school and teachers' college enrollments have increased from 35,000 to 162,000 and are estimated to reach 175,000 in 1940.

(5) Since 1890 the enrollments in colleges, universities, and professional schools have increased from 158,000 to 924,000. The rate of growth shows an increase of approximately 60% over each preceding decade. The enrollments in the levels of higher education exclusive of normal schools and teachers' colleges will probably reach 1,250,000 in 1940. In all levels of higher education beyond the secondary schools, the enrollments in 1940 will probably approximate 1,500,000 students.

*Teachers in all levels of instruction.* The data for the several decades have been expressed in the nearest thousands as are the estimated figures for the several teaching levels in 1940.

The number of teachers, expressed in thousands, for the several decades and types of schools are given below:

The data on the number of teachers for the decades 1890 to 1930 and the predicted enrollments of each of the several levels of instruction may be summarized as follows:

(1) From 1890 to 1930 the number of teachers giving instruction in the kindergarten and elementary schools increased from 376,000 to 703,000, a rate of increase much higher than the

| *Teachers* | *1890* | *1900* | *1910* | *1920* | *1930* | *1940* |
|---|---|---|---|---|---|---|
| Elementary, including Kindergarten | 376 | 429 | 516 | 621 | 703 | *750* |
| Secondary | 16 | 30 | 52 | 117 | 238 | *300* |
| Normal Schools and Teachers Colleges | .. | 4 | 6 | 9 | 14 | *16* |
| Colleges, Universities, and Professional Schools | 13 | 24 | 35 | 45 | 70 | *85* |
| All teachers | 417 | 504 | 631 | 815 | 1,038 | *1,200* |

rate of increase in school enrollments in these levels. The estimated number of teachers for this level of instruction in 1940 is 750,000.

(2) From 1890 to 1930 the number of secondary school teachers increased from 16,000 to 238,000, a rate of increase which kept pace with that of the secondary school enrollment. The predicted number of secondary school teachers for 1940 is 300,000.

(3) From 1900 to 1930 the number of teachers in normal schools and teachers' colleges increased from 4,000 to 14,000. The prediction for teachers in this level and type of instruction is 16,000 in 1940.

(4) From 1890 to 1930 the number of teachers in colleges, universities, and professional schools increased from 13,000 to 70,000, at a rate which almost kept pace with the five-fold increase in enrollments in these institutions of higher learning during the past four decades. The predicted number of teachers in institutions of higher learning for 1940 is 85,000.

Briefly stated, the instruction of students of all levels of education was done in 1890 for 15,000,000 students by 417,000 teachers; and in 1930 for 30,000,000 students by 1,038,000 teachers. It is here estimated that the task of instructing 36,000,000 students estimated for 1940 in the several levels of education will be carried on by 1,200,000 teachers.

*Preparation of Teaching Staff.* Teachers who enter the profession in 1940 will hold credentials showing educational training of four or more years beyond the secondary school. For teachers of elementary schools, both rural and urban, this period of training will include a curriculum of four years, approximately 60 per cent of the time will be devoted to the mastery of the humanities, the social sciences, the physical sciences, the biological sciences, English language and literature, practical and fine arts, with electives in the fields of the classical and modern languages, and in mathematics. That forty per cent of the four-year program of the elementary school teacher which is devoted to professional training will certainly include history and philosophy of education, educational psychology and measurements, educational methods and practice teaching. The four-year teachers' college program of training for elementary teachers will be more academic than now, while the training programs for these teachers as given in college and university departments of education will give more emphasis to classroom methods and techniques and to the theory and practice of teaching.

The elementary school teacher who enters the profession in 1940 will be equipped as a four-year college graduate with professional training. Elementary school teachers of 1940 with less academic and professional training than that described above will be rapidly displaced unless they continue their education to a point which approximates the four-year college program.

In our nation today but one state (California) requires of secondary school teachers one full year of graduate study, 33 states require college graduation, and 15 states do not even require graduation from college in the equipment of those who qualify for teaching in secondary schools. The trend toward college graduation (100,-

000 students are now graduated annually) with the presence of schools of education in our leading universities is certain to lead state boards of education throughout the nation to decide upon a nearly uniform practice of fixing five years of academic and professional training beyond the secondary school as the minimum requirement of secondary school teachers of academic subjects.

By 1940 the junior colleges, now numbering 450 and enrolling 75,000 students, will have reached an enrollment of 125,000 and will demand a type of teacher as well prepared in academic subject matter as is required for lower-division instructors in the college and university. For these schools, in their academic departments, there is certain in 1940 to be a demand for teachers who have professional training for teaching and at least two years of graduate study. The doctor of philosophy and the doctor of education will be found with increasing frequency in the ranks of the junior college teachers.

Educational literature in the college fields now indicates that graduate departments of some of our leading universities are interesting themselves in the preparation of their doctor's candidates for college teaching. Content is stressed; yet method is considered, though it seems to be agreed that the proper methods can be revealed and imparted only by Ph. D's with majors in subject fields. Method is here recognized as important in the equipment of the college teacher and a real victory has been won; by 1940 college teachers will have learned many of the applications of educational philosophy and educational science to college teaching. Our college teachers now at work will attempt to teach themselves improved methods while the on-coming generation of college teachers will secure this knowledge and training as a part of their graduate study and training.

Soon after 1940, all elementary teachers will have had four years of college training, the typical secondary school teacher will have a full year of graduate study to his credit; the junior college teacher, two years of graduate study; and the progressive college teacher will have had training in educational methods appropriate to his subject and directed practice teaching as an integral part of his preparation for teaching.

*Purposeful Education.* The teachers of yesterday taught the content included within the covers of the adopted texts. So do a majority of the teachers of today. The thought of one author in a given field almost circumscribes the thinking of the pupil. Learning is yet too largely a mere mastery of fact, with little consideration of the purpose, or plan, or method.

Fortunate it is that textbook writers are beginning to give consideration in their texts to the purposes to be achieved or goals to be reached by the pupil through the use of the printed pages. A beginning only has been made. Of the purpose to be achieved by his course, not one teacher in a hundred has a clear vision. Such progress as we have made is due largely to the work of Snedden, Bobbitt, Charters, and a few others. These men have forced certain teachers to face the question "Why do I teach Caesar, modern history, geometry, English literature, and the like?"

It is the opinion of the writer that the major purposes of the course should be known to the teacher, to guide in the instruction process, and to the pupil to guide him in his learning efforts. In 1940 to a larger extent than now, the mastery of content, the directed experiences of the student all will contribute to the achievement of the known and accepted goals of the course, which shall mark mile posts passed in the development of the student.

By 1940 our teachers will have moved far from the teaching of mere subjects. They will have accepted it as their responsibility to teach through directing student experiences with useful content. Teachers will realize more than now that mastery of subject matter is not an end in itself, but simply a means of bringing about desired ends, clearly set forth in such language as will be understood by teacher and student alike. The learning process will be participation in a series of activities in which each step has a meaning and contributes to the attainment of known and accepted objectives.

*School Classrooms in 1940.* Though notable departures may be seen in certain school systems, still the typical American school classroom has from 30 to 40 seats of the conventional type, the teacher's desk, the blackboard, the dictionary, the nation's flag, and little else. Certain writers in educational magazines are now pointing out the fact that real learning consists in responding to stimuli under the direction of the teacher. Here the teacher shows his skill and resourcefulness in arranging the elements in the school environment in which the pupil lives and learns by responding to stimuli. Obviously the elements which together constitute the environment of the classroom should be selected with a view to stimulating the type of response which is desired. Teachers of the sciences, and of the fine and practical arts have pointed the way. More recently teachers of classical and modern languages have seen the advantage of creating the classroom environment which provokes thought appropriate to the subject taught. From several quarters comes evidence of attempts to select and arrange the classroom equipment — its bulletin boards, its furnishings, its bookshelves, its draperies, its pictures, its blackboards, its chairs—in an effort to create environment appropriate to the learning activities which go on within the walls of the room. The old type of bare-walled classroom is sure to go by 1940. By then we will have made much progress in arranging our English classrooms for English thinking; our history classrooms for developing history mindedness; our mathematics classrooms for thinking in quantitative terms with mathematical symbols.

*Methods of Learning.* The efforts of teachers since the beginnings of formal education have been centered on the problem of pupil achievement. The "Why?" of the process has only recently come to educational consciousness, yet already gains in the direction of purposeful education have been observed as has been previously mentioned. Somehow knowledge has been held sacred and the gaining of it has been accepted as worthy of supreme effort. Its possession has been the desired goal. The processes by which

the goal is to be reached or has been reached are given a minimum of consideration or neglected altogether. The possession of knowledge has been and still is the all-important end.

As the writer views the matter, much more importance should be attached to the nature of the processes by which information is acquired and the problems solved. The particular fact may never occur again, and the problem in its present form may never reappear, yet other facts must be acquired and other problems must be solved. The mastery of method of work, therefore, takes on increased significance. The plan of work should be ever in the student's consciousness. From day to day, from month to month, and from year to year, the student should ever be alert in carrying his plans to their fulfillment and in improving his techniques of learning.

With our acceptance of student development as the goal of education, we are certain during the next decade to focus our attention on the methods employed by the student in gaining a real command of himself and in realizing to the full his maximum personal development.

In this connection, President von KleinSmid and I have collaborated in organizing and integrating the contributions of approximately one thousand junior and senior college teachers as these contributions have revealed important steps in the methods of work to be used by students in their attempts at the mastery of content elements in some twenty major fields of study in the junior college level of achievement. Since the content learned in school is so soon or so nearly forgotten, the justification of the effort expended must be in the mastery of methods used, rather than in the value of the facts learned.

By 1940, we will, I think, be taking this new view of the educational process. We will count our efforts worthy in proportion to the success of our efforts in developing student abilities and individual student methods in mastery of situations and problems, at least some elements of which are new or unfamiliar to the student.

*Use of Moving Pictures in Instruction.* Whether we wish to do so or not we are forced to recognize the moving picture as a powerful educational agent. It may give incorrect information and may set up wrong standards of life, yet children and adults do "go to the movies," and get while there a certain type and amount of education.

Visual aides are provided now in modern schools. As yet they supplement the normal procedures rather than supplant them. In the laboratory and lecture room the moving picture presents in fifteen minutes the visual record of the life history of the plant or animal; in the shop the details of the processes of industry are revealed by the moving picture records; the football and track coach uses the moving picture to point out the elements of good form and to check the accuracy of field performance.

The secondary schools will in 1940 make a large use of films to give visual pictures of scenes of literature and history, to vitalize project work of many types, and to exemplify types of activity where standardization is desired.

*Student Activity in the Classroom.* From the beginnings of formal education to the present, approved classroom procedure has been that of reciting by the pupil with the checking and correcting of errors by the teacher. Dur-

ing the decade just passed, and in the elementary school in particular, daring teachers have had the courage to vary the usual procedure so as to give time to the direction of pupil activity in new fields of work. This plan is finding its way into the junior high school, but has not been accepted as an approved method in the senior high school or junior college. By 1940, I think, we shall find many teachers in the secondary schools who will plan to divide the recitation period about equally between the checking of students' work and the directing of the thinking of students as they approach their new assignments. Of course, checking of work will continue to be important, yet a clearer understanding of the responsibility and opportunities of the teacher will increasingly draw the attention of the teacher to the richer field of effort,—that of directing the thinking of students as they undertake their new tasks. The longer or sixty-minute recitation periods in some secondary schools make it possible for the teachers to give some attention to the review of assigned materials and yet leave sufficient time to observe and direct pupils as they initiate their new tasks. The secondary school classroom of 1940 will be one in which work is accomplished with a minimum of attention given over to the checking process.

The college class exercise seems doomed to change also. The lecture method is almost certain to be replaced by a lecture-conference method in which a maximum of student activity is evident. A syllabus will be worked out for each of the college courses. In this syllabus problems will be raised and bibliographical references will be given. Students will read several from among the many references given. The class period will then be devoted to the processes of integrating the opinions or pieces of objective evidence which are presented by the several authors to whom references were made. In this way students will be trained in sifting evidence and in organizing materials judged pertinent to the issue under consideration.

*Equality of Educational Opportunity.* Obvious inequalities in educational opportunities are found today throughout our nation; inequalities almost as great are to be found within the borders of many states, and even within the boundaries of the counties within a given state. Local option in education has been taken to mean that each school district is privileged to have as poor a school as it wishes. Also, many school districts vote taxes heavier than common property can bear.

The writer is unable to predict what the sources of school income will be; that is, whether real estate and mortgages shall constitute the bulk of taxable wealth or whether income taxes, inheritance taxes, sales taxes, and intangibles will have been shown to be the better sources of revenue. The fact with which the writer is chiefly concerned is that the wealth of the nation and state shall be taxed to support the education of the children of the nation and state. Wealth must be taxed where it is and revenue thus gained must be spent where the children are. Inequalities in educational opportunity must be ironed out; all children must be given equal opportunities for education through the secondary school period. Students enrolled in levels of education beyond the secondary school may be called

upon to bear from one-third to one-half of the costs to the taxpayer which are entailed in the instruction and other educational provisions which they use. It may be shown in 1940 that free tax-supported education beyond the secondary school shall be limited to a few who are exceptionally able and who need this subsidy. Others may be required to pay tuition covering approximately one-half of the instructional costs.

The facts on mobility of the population of our nation—from one state to another, from the country to the city, from one city to another — all argue for good educational opportunities for children in the areas where they are born. It now makes a great deal of difference to Long Beach, California, whether or not the educational opportunities are good in the rural areas of Iowa. In California today there are 3,600 school districts with as many possible varieties in the adequacy of school support.

We recognize our small district system as a Massachusetts inheritance. The plan was used there but soon discarded. Perhaps the nation will have made progress toward larger school districts in 1940. According to the provisions of our National Constitution, education is a business allotted to the state for control. Would that the state would assume the responsibility of directing the education of the children to the end that all children in the state may have adequate and nearly equal educational opportunities, whether the child must live in the city, in the mountains or desert, or in agricultural areas. Our present plan of control and support of education is nothing short of opportunism and political expediency; it smacks not of clear thinking or sane judgment. It is the prediction of the writer that we will have made considerable progress toward state support of education and equality of educational opportunity by 1940.

*Conclusion.* It is predicted above that the enrollments in all levels of Education in 1940 will be as follows: Kindergarten and Elementary, 26,000,000; Secondary, 7,000,000; Normal Schools and Teachers' Colleges, 175,000; Colleges, Universities and professional schools, 1,250,000; and Special Schools, 700,000. It is estimated that the teaching of the 36,000,000 students of 1940 will be carried on by 1,200,000 teachers. Soon after 1940, all elementary teachers will have four years of college training, while secondary teachers will have five years, and junior college teachers will have six years of college training. Moreover progressive college teachers will have had training in educational methods. Teaching in 1940 will consist in directing student experiences toward clearly defined objectives with appropriate content. Classrooms will be built and equipped with special reference to subjects taught within their walls. Methods of learning will consist of a conscious well-directed attempt to develop abilities of individual students, films will be used to present scenes in literature and history, and processes in physical and biological sciences. Students will be trained in organizing materials and making judgments in classroom activities. A clearer vision of the need for education in our democracy will lead in the direction of state-supported education and equalized educational opportunities through the secondary school period.

# THE FUNCTION OF EXAMINATIONS

MARK E. HUTCHINSON

DEPARTMENT OF LATIN
CORNELL COLLEGE, MOUNT VERNON, IOWA

*The relationship between measurement and marks is so close that to most college students the two words mean the same thing. The chance of securing employment, the opportunity for promotion, as well as the feeling of self-satisfaction are all governed very largely by the mark that one receives from his instructors.*

IT seems to me that examinations should accomplish at least three things: (1) they should help the teacher form an accurate judgment of the students' attainments, (2) they should help the student organize his knowledge, if he has any to organize, (3) they should give aid to both student and teacher in diagnosing the weaknesses and lacunae in the student's knowledge. By examination, I am referring to an achievement test in some college subject. As a matter of fact, many of the examinations which we give fall far short of attaining these objectives, which are not so easy as they sound. Objective number 1 is particularly elusive, "to help the teacher form an accurate judgment of the students' attainments." Many teachers have their minds made up before they give the examination what the student will do on it and this influences their judgment which is expressed by the grade or mark. It is the mark which causes all the trouble and around which a battle in educational circles is raging this very hour. Of course credits and marks are inseparable and they are the chief cornerstones of our educational structure. Many teachers and administrators are beginning to question our present credit and marking system and "that mogul of modern education," the college registrar, is tottering on his throne or perhaps better his swivel chair. Professor Henry Crew in his presidential address given before the 1929 meeting of the American Association of University Professors says: "This same principle, when applied to our marking system, brings out with frightful clearness the horrible mechanism by which we are trying to provide mass education and keep the lame ducks in the procession. Our marking systems, whether of daily, weekly or monthly grades, divert the student's attention from the main thing, produce a mistrust in his mind, give him an "on guard" attitude towards his professor, and lead him to conceal his ignorance, instead of indulging his natural and divine curiosity in frank discourse with his preceptor. But this is only half the story of marks. The instructor who examines on the *subject* rather than on the *text* — the teacher who employs an attainment test rather than a trial of the memory —is under the present machinery almost certain to suffer one of two evil consequences: either he reports the "standard" percentage of his students as passed and loses his self-respect, or he reports the facts and fails so many students that his dean upbraids him for poor teaching."

Let us put the cards on the table. We all know that in all of our colleges three hours under Professor A are

worth six hours under Professor B, so far as student effort and grades are concerned. We know perfectly well that an A from Professor X is worth a C from Professor Y. But what can we do about it? We steadfastly maintain that our own system of marking is above reproach, or practically so, but when it comes to our colleagues we could find much to criticize, if we were so inclined. In other words, credits are based on marks, and the marks are far from accurate. Even if they were accurate, too many students are interested in grades rather than in accomplishment. Nor do I blame them so much as the system. Percy Marks in his "Which Way Parnassus" may exaggerate a little, but very little, when he reports the following outbursts from students: "I worked like the devil in Math. and only got a D. The folks were awfully sore. I did all of Jim Jones' problems for him and he got a B. Believe me I am not going to be a fish again. Watch me get an A." Another outburst, "I am sick of busting my neck over these blamed English themes. I sat up half the night working on one and only pulled a D. And Bill Smith copied his out of the Lit. Dig. and got a B+. You can bet I'm not going to lose sleep any more just for nothing. I am going to buy the Lit. Dig. myself." Mr. Marks says the most honest examination ever handed in to him was as follows: "There's a beautiful fog down on the beach today, and I don't feel like playing at being Mr. Wiseman with you. You can play at being Mr. Wiseman with the other little boys. I am going down and play in the fog."

Dean Haggerty of the University of Minnesota has said: "There can be little effective study of college teaching and little intentional improvement in it until we discard the present marking system, or supplement it with corrective devices that render it intelligible. More than twenty years ago it was clearly demonstrated that the marking system employed in most schools and colleges is statistically unreliable. Since that time scores of statistical studies have been made all confirming this fact. Notwithstanding the accumulating evidence, colleges have been slow to devise better methods and have for the most part declined to adopt the improved techniques recommended to them. . . . Investigation has not revealed any complete panacea for the defects of the present marking systems, but two methods of improvement that promise usefulness may be commended. The first is the percentile technique of recording student standings in examinations and courses: the second is objective examinations."[1]

At the University of Chicago an experiment has been started where credits and years will be disregarded and *achievement* will be measured, as accurately as possible. The measurement by time will be done away with and the students' achievement will be measured by a comprehensive series of tests, many of them objective, constructed by both subject specialists and experts in the testing field. Reports coming from the University of Chicago indicate that this new plan of scrapping the old system of marks and credits has made a good start. At Oberlin College beginning with the year 1930-31 there was adopted a

1 Cf. Raymond A. Kent, "Higher Education in America": Ginn and Co. (1930), 507-508.

ranking system instead of the usual method of giving literal grades. At the end of the semester each teacher sends to the registrar's office a list of his students in the order of excellence. The ranks of the passing students are then converted into a score by the registrar by means of a conversion table.[2]

All of this may seem rather far afield from my subject, "the function of an examination," but there can be no discussion of examinations without a criticism of the system of grades and credits now generally in use. Surely our examinations test us who make them as much or more than they test those who take them. Several years ago 116 candidates for teaching certificates in the state of New Jersey took an examination and not one of them passed it. As a result, the State Commissioner of Education ruled that this examination did not constitute "an adequate instrument of the work of the normal schools and can not be regarded as a valid test for the selection of teachers." And that brings us to the most important thing we can ask ourselves about our examinations—are they valid? That is, do our examinations test for what they are supposed to test. A test on Art Appreciation should test for art appreciation and not for the ability to draw or paint or what not. In order to make a valid test an instructor must know what he wishes to accomplish in his course and his students must be made aware of his aim. A teacher who has badly organized his course or does not know his subject will not make a valid test. In his efforts to make a valid test the instructor will construct his course in orderly fashion. He will have certain objectives and his teacher procedure will be directed towards attaining these aims.

There can be little doubt that objective tests are more likely to be valid than the traditional examinations in essay form. Therefore all college teachers should experiment with this type of test and become familiar with the technique of constructing such tests. Not only should teachers construct objective tests for use in their teaching, but they should use standardized tests in their own field, so far as they are available. Of course the chief purpose of a standardized test is to compare what your students are doing with what the average or norm is accomplishing and then to bring your students up to the standard or beyond, if humanly possible. The instructor can also discover certain habitual errors which his students are making. While I believe that a well-made objective test is more likely to be valid than an essay form of test, we should use both kinds of tests, for after all the student needs some practice in writing English and must not spend all of his time checking true-false and multiple-choice exercises. If I were asked by some particularly credulous reader of this paper just how he should use these new-fangled tests, as he might call them, I would advise him (1) to give his students a placement examination in order that they may be placed with students of approximately the same attainment.[3] (2) to give early in the course a good standardized achievement test in order to

2 Cf. Carl C. W. Nicol, "The Ranking System," The Journal of Higher Education, III (1932), 21-25.

3 The Iowa Placement Examinations and Aptitude Tests are probably the best known tests of this nature. They can be obtained from the Bureau of Educational Measurements, State University of Iowa, Iowa City, Iowa.

see what his students actually know (3) to give a duplicate form of this same test towards the end of the semester so that he may find out whether they have learned anything. (4) to make his own informal objective tests and make them as valid as he possibly can. These should be given frequently and they will serve as teaching devices as well as tests. (5) to give a comprehensive test to all of his major students before they graduate. The nature of this test would vary with the department.

The more one works with tests and examinations the more fascinated he becomes with the subject. Doctor F. J. Kelly has laid down the following as needed research along the line of examinations and marks. I see no reason why any self-respecting college should hesitate to tackle these problems. They are: "(1) a statistical study of marks given by teachers. They should include (a) distribution by teacher, department and school (b) correlation of marks in various subjects with other criteria of ability such as intelligence tests, vocational success or part taken in worthy non-vocational enterprises. (c) relation between marks and withdrawal from college. (2) experimentation with the development and use of the more objective forms of tests and examinations. (3) experimentation with substituting for grades and credits in the individual subjects in the senior college a single comprehensive examination over the whole field of concentration."[4]

The moral of this paper, if it has one, runs somewhat as follows: Our method of marking and giving credits is faulty. We are too much the devotees of time and the registrar's office. The way out is a difficult one and seems well concealed but we must find it. Perhaps Chicago is blazing the trail. Perhaps not. We should be sure that our examinations are valid. Our students should be examined soon after entrance by placement and achievement tests and placed where they belong regardless of the amount of time they have or have not spent on the subject. We should consider seriously the advisability of introducing comprehensive examinations for our major students or perhaps for each field of concentration. We should not think our examinations perfect, even if they be objective tests. We should possibly adopt some method of percentile ranking of our students rather than giving them an A or flunking them. Until this method is adopted officially, we can use it unofficially in our own classes.

Finally, when we give our examinations let us avoid the following objectives with which, according to Dr. C. C. Little, the former President of the University of Michigan, most oral examinations for the Ph. D. degree are given. These objectives are: "(1) to impress the other examiners with the fact that he, the questioner, has kept up to date intellectually. (2) to bewilder the candidate by asking questions which he cannot answer. (3) to hide the fact that the examiner is unfamiliar with the field in which the candidate is specializing."[5]

[4] Cf. F. J. Kelly, "The American College": New York, Macmillan Company (1925), page 133.
[5] Cf. C. C. Little, "The Awakening College": W. W. Norton and Co. (1930), page 136.

# BREAKING THE FOREIGN LANGUAGE GOOSE STEP*

WALTER VINCENT KAULFERS

LONG BEACH JUNIOR COLLEGE, LONG BEACH, CALIFORNIA

***A plan that "leads to the acquisition of measureable skills"; that "makes measureable achievement the basis of the marking system"; that "increases enrollment in the upper divisions"; and that "lessens the instructional labor of correction," sounds almost too good to be true.***

IN a sense, the *goose step* in the foreign languages constitutes far more a pragmatic ideal than a necessary evil. If the ordinary language pedagogue could align his charges into companies of identical size and stature, and could make all advance with equal pace toward the simultaneous achievement of a uniform goal, he would not fail to think that the *summum bonum* and the *ne plus ultra* in class organization had been accomplished.

That a mechanically controlled homogeneity of achievement in the foreign languages is next to impossible, however, scarcely requires confirmation. Every teacher of experience has found in his beginning classes pupils who at the end of the term acquired as much ability in vocabulary, oral reading, or comprehension, as students in the third and even fourth semesters — pupils who, had the opportunity been afforded them, might easily have made similar progress in grammar and composition. Conversely, there is no instructor of foreign languages who has not entertained in his advanced courses students who from all the evidence of standardized tests had never progressed beyond the elementary stages. The prevalence of this condition in the foreign language field has been made patent by the comprehensive investigations of Ben D. Wood in New York, and by every study of the problem reported since that time.[1] To illustrate the gravity of the situation a few examples may be taken from life. The following are actual cases discovered in a study of 663 pupils of beginning Spanish in eighteen junior high schools of Los Angeles and Long Beach in 1931:

*Case A,* a boy enrolled in ninth grade beginning Spanish, Long Beach, IQ 91, made a total score of 185 on the *Columbia Research Bureau Spanish Test.*

*Case B,* a girl, IQ 110, enrolled with the same teacher, made a total score of 23 on the same test.

*Case C,* a girl enrolled in beginning eighth grade Spanish, Los Angeles, IQ 118, made a total score of 0 (zero)!

*Case D,* a girl of the same IQ, enrolled in the same grade, achieved a total score of 229!

These examples, though extreme, are indicative of a general condition.

The challenge of the problem has been met in some institutions through differentiated assignments, supplementary readings, and honors courses, but to the present time the principles governing the organization of a remedial plan, and the mechanics of its administration, have not been sufficiently

* The essential features of the plan have long been in successful operation in the English, Mathematics, Science, and Social Studies departments of the Woodrow Wilson High School, Long Beach, California, under the direction of Dr. William J. Klopp, director of teaching.

1 Ben D. Wood. "New York Experiments with New Type Modern Language Tests." The Macmillan Company, New York, 1927, 339 pages. See also: Walter Kaulfers. "Effect of Mental Age on Foreign Language Achievement," in Journal of Applied Psychology, vol. XI, no. 3, pp. 257-268 (June, 1930).

understood to lead to a concerted program of reform. Although it is still impossible to point to such a solution in practice as promises feasibility for all levels of foreign language work from the junior high school to the college, sufficient experience in other academic fields is available to furnish the basis for a mode of attack. At least the following premises may be considered corroborated by scientific evidence:

1. No solution is likely to prove efficacious which depends upon a homogeneous grouping of pupils into fast and slow moving sections, for not a single basis of classification, from IQ's and prognosis tests to measures of vernacular ability in English and General Language, yields a sufficiently high correlation with achievement, either singly or in combination, to assure an accurate individual placement of beginners for all schools.[2]

2. In general, the problem constitutes as much a question of individual differences in interest, motivation, and industry as of aptitude or capacity, and in consequence cannot be solved solely by prognostic methods. It compels such a motivated individualization of instruction as makes possible for each student a rate of progress commensurate with his interests, abilities, and ambitions.

The only obstacle in the way of such an ideal is the feasibility of its application in school life. To be at all expedient in practice, the plan requires the organization of subject matter into a graduated series of specific content-units, or mastery-unit exercises, suitable for successive completion by the student independently of the class and of the instructor. Obviously, the exercises must be sufficiently limited in scope to permit of brief explanations for self-instruction, plentiful enough in practice material to insure a mastery of the principles or content involved, and sufficiently simple and objective in organization to permit of ready correction from uniform keys.[3] Beyond this it is desirable that norms of achievement be reported on each unit for the edification of the student in orientating him in standards of work, as well as for the information of the corrector in the subsequent scoring and grading of papers. The norms can in most cases be determined through preliminary tryouts on representative class groups, and can be revised or extended with experience from year to year. For those students who fail in attaining the minimum standard, alternate exercises are essential for supplementary drill. Finally, a set of brief 10 to 15 minute objective tests—one test of two forms for each unit — with corresponding norms, is imperative. For facility in administration, the material should be mimeographed and bound in booklet form, the syllabus of mastery-unit lessons to be kept by the pupil, and the pamphlet of tests by the instructor.

When the student shows evidence of having completed satisfactorily a unit of work, he applies for the corresponding mastery-unit test. On the designated examination days he receives his quizz-book, and demonstrates his ability to apply the subject matter of the lessons independently of outside aid. The *order* in which the mastery-unit tests and subsequent practice exercises are completed is uniform for all stu-

2 Walter Vincent Kaulfers, "Present Status of Prognosis in Foreign Language," in School Review, vol. XXXIX, no. 8, pp. 585-596 (October, 1931).

3 For a complete discussion of the construction and validation of mastery-unit exercises, with practical illustrations for a specific language, see: Walter V. Kaulfers. "The Mastery-Unit Plan in Foreign Language Teaching," in Hispania, vol. XIII, no. 5, pp. 417-430 (November, 1930).

dents, but beyond the motivation necessary in special cases, the *rate* at which the units are finished is left to the pupil.

As soon as all units for the semester have been mastered, and in any case at the end of the term, the student is given a comprehensive standardized test to determine his general achievement in the language. If his score attains the norm for the course, he is given full credit for the semester, and, provided the class schedule permits of transfer, is promoted at once to an advanced class;[4] otherwise, he begins work on the succeeding units while remaining with his group.

Marks under this plan are awarded on the basis of both rate of progress and proficiency in subject matter, as follows:

A grade of *A* (excellent) is earned by

1. Achieving a score on the comprehensive examination within ± .5 standard deviation of the norm at least seven weeks before the end of the term; or by
2. Achieving a score at least + 1.5 standard deviation above the norm at the *end* of the term.[5]

A grade of *B* (good) is earned by

1. Achieving a score on the comprehensive examination within + .5 standard deviation of the norm at least four weeks before the end of the term; or by
2. Achieving a score at least + .5 standard deviation above the norm at the end of the term.

A grade of *C* (fair) is earned by achieving a score on the comprehensive examination within ± .5 standard deviation of the norm at the end of the term.

A grade of *D* (passing) is earned by achieving a score between −.5 and −1.5 below the norm for the comprehensive examination at the end of the term.

No grades below *D* are given, since sub-average achievement is adequately penalized by a reduction of credit.

For purposes of class instruction, the recitation time in fifty-five minute periods is budgeted as follows:

The first fifteen minutes on Tuesdays and Fridays are devoted to the administration of mastery-unit tests. Although relatively few students are eligible for the same examinations on the same days, no confusion arises from this source, since the questions are planned to occupy the full time of each pupil during the testing period, and the time-limit in all cases is uniform. At the beginning of the hour, the mimeographed quizz-books (previously collected in order of seating to facilitate distribution) are passed to the pupils. The desired exercises are readily accessible, for all completed tests have been torn out, thus leaving the appropriate questions exposed when the cover is opened.

The *first* fifteen minutes on Mondays, Wednesdays, and Thursdays, and the *second* fifteen minutes on Tuesdays and Fridays (the testing days), are reserved for the open-forum discussion of language problems raised by the students. Typical questions for this period are the position or agreement of adjectives, the distinction be-

4 In answer to the objection that this plan runs counter to university requirements, it should be pointed out that since the problem is ultimately one of bookkeeping, and affects only a small percentage of the students, the difficulty can be met in most cases by enabling pupils who transfer from a lower course to a higher to receive credit in both classes—i. e., the transcripts may show simultaneous enrolment in two courses.

5 Scores more than −.5 standard deviation below the norm on advanced examinations are disregarded in grading, since the students are obliged to take another form of the test at a later time before marks and promotions are granted.

tween the imperfect and the preterite, the uses of the subjunctive, and similar items, depending upon the nature of the course. These problems are discussed by the instructor before the class, in such a way as to capitalize the knowledge of pupils who have already completed the corresponding units of work. For the precocious students this period serves as a motivated review, while for the laggards it affords an invaluable opportunity for acquiring advance information on vital questions that all are destined sooner or later to meet. It cannot be urged too strongly that instruction given in these circumstances, in response to felt needs, is destined to be attended far more intently by the group, and therefore to be fixed more permanently for subsequent application, than information offered gratis on less auspicious occasions.

The remainder of the class hour—forty minutes on three days, twenty-five minutes on two days—is devoted at the instructor's discretion to class recitation, or, in junior and senior high school, to supervised study with special aid on personal problems. In order to allow as much time as possible for individual progress, no uniform assignments are given. In other words, the group work is limited to supervised sight-reading from the text (with questions and answers); informal dramatizations; games or songs; and discussions by the instructor of selected phases of foreign life, literature, and custom. Reading is chosen as the major class activity since it comes nearest toward representing the ultimate aim of foreign language instruction in the United States,[6] and because in groups of varying individual abilities and preparation, such as must be anticipated under the plan, it is perhaps the only type of class work that permits of a profitable participation on the part of all pupils. To avoid the repetition of class work for students who fail of promotion to advanced groups, a different reader from the one used in the fall is read in the spring.

Inasmuch as the difficulties surrounding such a proposal — the want of suitable instructional materials,[7] the disagreement among teachers concerning aims and standards of achievement, etc.—are too obvious to require mention, the discussion of the project may more profitably conclude with a consideration of possible merits. In favor of the plan the following advantages are proffered:

1. It shifts the emphasis from *time-serving* to *achievement* in the adjudication of credit, thus making each semester of foreign language work lead to the acquisition of measurable skills and abilities, rather than to the mere accumulation of credit hours.

2. It makes *measurable* achievement the basis of the marking system, thus stimulating a higher type of scholarship by reducing the popularity of the "gentleman's C" and the "bench-warmer's D."

3. It enables the best students to attain the advanced stages of foreign language work in from two-thirds to one-half the usual time, thus operat-

6 Algernon Coleman. "The Teaching of Modern Foreign Languages in the United States." The Macmillan Company, New York, 1929, 299 pages; page 170.

7 In large schools the composition of the units can be made a departmental project, each teacher undertaking the treatment of a specific topic. This division of labor materially reduces the load per teacher and eventually results in a comprehensive library of useful material, while the subsequent pooling of exercises makes possible interesting comparisons between different groups. If the exercises are made self-instructional, they may entirely displace the usual grammar.

ing effectively to increase the enrollment in upper division classes.

4. It is adapted alike to intra-group and self-competition, thus exercising a salutary influence upon achievement and scholarship.

5. It lessens the instructional labor of correction: the objectivity of the tests and exercises permits of a ready checking and scoring of responses from uniform keys by advanced students.

6. Since the number of promotions within any semester amounts only to a limited percent, the heterogeneity of class groups becomes no greater under the mastery-unit plan than under present conditions; while with reading as the major class activity, individual differences can readily be accommodated in the recitation.

7. The plan requires a scientific organization of subject matter into psychologically spaced units of gradually increasing difficulty, with norms and measures of achievement. Consequently, it stimulates a wholesome analysis and re-evaluation of content, activities, and standards on the part of foreign language departments.

## YOUR SCHOOLS AND MINE

ROY C. WOODS, PH. D.
PROFESSOR OF EDUCATION, MARSHALL COLLEGE, HUNTINGTON, WEST VIRGINIA

***In this department Dr. Woods will answer to the best of his ability any or all questions on educational topics that are addressed either to him at the address above, or to EDUCATION. It is to be understood that the writer alone is responsible for the answers given.*—H. B.**

*"My dear Dr. Woods: May a reader take exceptions to what you say as well as ask questions? If so, I want to say that I think the teacher you quoted in the December issue in her solution of the problem of snobbishness was to be condemned instead of held up for commendation. Aside from the pragmatic consideration of "having worked in this case," it has nothing to commend it. It resembles the "eye for an eye" doctrine too much to suit me. I think she should have had talks on sportsmanship, democracy, rights of the other fellow, etc. She should have used the class period to teach proper citizenship, school spirit and the like. This seems to me to be the positive approach, while the other the negative one. . . ."*

I submit this excerpt to show to all two things. Yes! you make take exception and I am pleased to hear from you. Furthermore, I am going to say I agree with what you have said. The only trouble is you didn't say enough. What are you going to do when your method failed. I have learned, or think I have, that there is no sure cure for any set of problems we teachers meet. What works with one pupil will fail with another and will have opposite results, i. e., antagonize another. In the case I quoted, "it worked," as you say, and I think that "nothing succeeds like success." I should rather solve it your method, but when it fails then something more severe is needed.

May I illustrate with expulsion?

Ordinarily I object to the use of it in discipline. Legally no one but the board can expel a pupil. Teachers are only permitted to suspend. It closes the doors of opportunity and in many cases merely confirms the offender in his anti-social tendencies. But if all other means have failed and the offender still violates his privileges, suspension will at least protect other members of the group who are social in their outlook and actions. Many times, too, it "works."

May I hear from you again as to how you would handle the case if your above methods failed?

*What is the status of manuscript writing? Has it proven its worth in classroom methods?*

Personally I doubt it. I am inclined to feel that it is still in the experimental or opinion stages. Mossman, in her recent book published in 1929, "Teaching and Learning in the Elementary School," says that it "recently came to us from England. It is being used in a large number of private schools and some public schools. Study is being made of the results, so that, in time, definite data as to its value will be available." I suggest that you turn to page 200 of the above reference for her discussion of its values and limitations.

*I am the principal of an elementary school and would like to know when I should schedule my teachers meetings. Should they be at the noon hour or after school?*

As a principal you should read Reavis, Pierce and Stulken's new book "The Elementary School." You will find some valuable suggestions in the book relative to administration of this type of school. I frequently refer to it in my work. They suggest that two types of meetings be held. One, of a short administrative nature, at the noon hour any day, where routine matters can be dispensed with in short order. The longer one, of a scientific nature, to be held after school at a convenient date for the teachers, where longer discussions of policy and theory or practice can be held. You will find that no date will be satisfactory to all, yet you can arrange one that will suit most.

*Do you feel that the segregation of handicapped children into specialized rooms is desirable?*

In most cases, yes. Certainly no one would say that tubercular children, certain types of disciplinary cases should be kept in with normal children. Deaf-mutes, blind children and crippled ones will need part segregation, depending upon their handicap. Until they can profit more by ungrouped classes they should be segregated. As to ability grouping, which is probably what my questioner has in mind, you will find wide differences of opinion. I think such grouping, properly handled, to be distinctly advantageous. McGaughey is one of the most active opponents of the movement, while Terman, Otis, Thorndike, Freeman and others, defend its value.

# DEPARTMENT OF BOOK REVIEWS

**Introduction to the History of the Social Science in Schools.** By *Henry Johnson*, Professor of Education, Teachers College, Columbia University. Scribner's, 1932. 145 pp.

The broader the base relative to the height of an object, the more stable it becomes. This volume by Dr. Johnson broadens the basis in the field of social science, by reviewing the important methods of teaching which have been used or advocated through the centuries. It is illuminating to learn that Weise as early as 1680 emphasized the vital unity of geography and history; that Basedow, 1770, advocated teaching history backward, beginning with the present problems and the use of national holidays as centers of interest for backward looks; that Christian Trapp about 1800, introduced notebooks much like our present work-books and also used a form of true-false test; that loading the minds of children with poorly comprehended facts was condemned by Frederick the Great, who is entitled to be ranked as an educator as well as a ruler and general.

The present combining of subjects into the Social Sciences may be traced from Pestilozzi through Harnich, Ziller's treatise on Herbart, and the American Herbartians, De Carmo, Frank and Charles McMurry, and Francis W. Parker. The idea of history as a means of propaganda or indoctrination was alive at the time of the French Revolution and has been since used in all Central European countries. Thus we discover that the Social Studies Report of 1916 contributes scarcely a new idea.

Dr. Johnson's treatment is a brief one, reads easily, is pointed and definite in its viewpoint. It shows the origins of present practices. Any teacher of history or the social sciences would be greatly stimulated by reading it. It gives perspective and makes present workers realize that creative ideas in the field, however new they may appear to the creator, have in most cases been "new creations" again and again through the centuries.—GUY M. WILSON, Boston University.

**Administration of the Testing Program.** By *Clifford Woody*, Director of Bureau of Educational Reference and Research, University of Michigan, and *Paul V. Sangren*, Director of Bureau of Educational Research, Western State Teachers College, Kalamazoo, Michigan. World Book Company, pp. 397, price $2.00.

In the Educational Review for March, 1926, Dr. Guy M. Wilson of Boston University called attention to the fact that the ends which should be served by testing were more fundamental than the testing itself; that a test that did not reinforce right curricular principles, encourage proper methods of teaching and serve the true purposes of an examination, i. e., provide for a new view or worth-while applications of material taught, was to be rejected regardless of its statistical and mechanical excellence.

In other words, the administration of the testing program is quite as important as the tests themselves. The authors have here answered the many questions regarding testing that have been asked them as directors of research bureaus. Out of their many years of experience in this field they have made a job analysis of the tasks involved in an effective testing program and have presented here a summary of the techniques involved in the administration of the program.

Such topics as scoring the tests, tabulating the results, using test results for the improvement of instruction and acquainting the public regarding the efficiency of the local school system, are treated in most interesting and practical fashion.

The last two chapters (Dangers to be Avoided in Using Tests and Present Tendencies in Measurement) should be made

required reading for all teachers and administrative officers. Each chapter would make a most excellent topic for discussion at a general teachers meeting.

**Modern Business Geography.** Revised edition, 1932. By *Ellsworth Huntington*, Research Associate in Geography, Yale University, and *Sumner W. Cushing*, late Head of the Department of Geography in the State Normal School, Salem, Massachusetts. Introduction by *Frank M. McMurry*, Teachers College, Columbia University, 352 pp. Price $1.96. World Book Company, Yonkers-on-Hudson, New York.

The authors are well fitted by training and experience for the piece of work undertaken. Dr. Huntington of Yale is an outstanding climatologist and a careful student. The late Sumner W. Cushing was an able young scientist and a normal school teacher of high order.

Geography was originally considered to be a subject for the upper schools. For many years, however, it was looked upon as being exclusively an elementary school subject—and it still is by many. Geography is again slowly but surely winning its way into the secondary schools. Unfortunately, high school geography has been associated with anaemic courses in comercial geography, in which stress has been placed upon memorization of lists or products and their sources. Little emphasis has been placed upon the "why" side of geography; and geography on any level is essentially a study of relationships. Huntington and Cushing's Modern Business Geography, while an economic geography, has not neglected this very important aspect. The student is called upon to consider problems for which the text has given the necessary background and to make use of geographical facts in thinking for himself. The book shows how the world's work is done and how it is affected by geographical conditions. This stress on cause and effect makes the book of real value in the teaching of geography.

Especially valuable are the problems and exercises found at the end of each chapter. They require thought to answer. Of course, without an enthusiastic teacher, even these problems may not function as intended.

One of the outstanding features of the book is that of making "figures talk." Making statistics meaningful is one of the outstanding contributions of the new school of geography teachers. The statistics are as up-to-date as it is possible to have them in a geography text-book.

There is a wealth of illustrative material and the pictures are well chosen. There is a splendid set of maps and graphs well adapted to a text-book in the field of economic geography.

The book is to be commended for its mechanical makeup, also. The convenient size of the book, the quality of the paper, and the clearness of details in the pictures, maps, and graphs, are worthy of mention. An excellent list of references in the field of business geography is given.

The Introduction by Professor Frank M. McMurry of Columbia University, while brief is a clear and keen statement as to what business geography should aim to achieve. The teacher will be repaid by making a careful study of this statement—made by one who has been out standing in the field of geography teaching for many years.—FRANKLIN C. ROBERTS.

**How to Make Pupil-Made Lantern Slides.** By *G. E. Hamilton*. Meadville, Pennsylvania: Keystone View Company, 1932, pp. 10.

Mr. Hamilton is the author of many practical articles and pamphlets on Visual Instruction. His latest one will be most useful in solving the problem of the home-made lantern slide, which has become so popular. It tells about the pencil-made lantern slides, ink-made lantern slides, Keymo lantern slides, and cellophane lantern slides. Teachers and pupils are obtaining excellent results with clear and etched glass illustrations of their own classroom activities. This booklet should certainly be in the possession of every teacher who is using visual aids.—F. D. McC.

# NEW EDUCATIONAL BOOKS

We acknowledge receipt of the following publications of interest to the readers of EDUCATION. Mention here does not preclude review in this or subsequent numbers of EDUCATION.

---

**Our Nation Begins.** By *Eugene C. Barker*, Head of the Department of History, University of Texas, *William E. Dodd*, Head of the Department of History, The University of Chicago, and *Walter P. Webb*, Associate Professor of History, University of Texas. 1932. 342 pp., illustrated, cloth $1.00. Row, Peterson & Co.

**The Story of Our Nation.** By the same authors as above. 1929. 401—xxii pp., illustrated, cloth $1.28. Row, Peterson & Company.

**Our Nation Grows Up.** By the same authors as above. 1932. 334-18 pp., illustrated, cloth $1.00. Row, Peterson & Company.

**Old Europe and Our Nation.** By *Eugene C. Barker*, Head of the Department of History, University of Texas, *Frederick Duncalf*, Professor of Medieval History, University of Texas, and *Francis Leonard Bacon*, Principal of Evanston Township High School, Evanston, Illinois.. 476 pp., illustrated, cloth $1.28. Row, Peterson & Company.

**The Growth of a Nation.** The United States of America. By *Eugene C. Barker*, Head of the Department of History, *Walter P. Webb*, Associate Professor of History, The University of Texas, and *William E. Dodd*, Head of the Department of History, The University of Chicago. 686—xxxvi pp., illustrated, cloth, $1.80. Row, Peterson and Company.

**The Human Personality.** By *Louis Berg*, M. D., Deputy Medical Supervisor, Department of Health, New York City, and Member of the Faculty, School of Education, Washington Square College, New York University. 321 pp., cloth, 1933. $3.00. Prentice-Hall, Inc.

**Lateral Dominance and Visual Fusion.** By *Charles A. Selzer*, Head of the Department of Education, State Teachers College, West Chester, Pa. The Harvard Monographs in Education, No. 12. 1933. 119 pp., paper, $1.00. Harvard University Press.

**The Evolving Common School.** The Inglis Lecture, 1933. By *Henry C. Morrison*, Professor of Education, The University of Chicago. 1933. 62 pp., cloth, $1.00. Harvard University Press.

**Pompilia and Her Poet.** By *Harriet Gaylor*. Third edition (Revised 1932). 198 pp., cloth. Modern Classics, publishers. $2.00.

**Educational Yearbook** of the International Institute of Teachers College, Columbia University, 1931. Edited by *I. L. Kandel*, M.A., Ph.D., Professor of Education and Associate, International Institute, Teachers College, Columbia University. 721 pp., cloth, 1932. Bureau of Publications, Teachers College, Columbia University.

**Administration of the Testing Program.** By *Clifford Woody*, Ph.D., Director of Bureau of Educational Reference and Research, University of Michigan, and *Paul V. Sangren*, Ph.D., Director of Bureau of Educational Measurement and Research, Western State Teachers College, Kalamazoo, Michigan. 1933. The World Book Company. 397 pp., cloth, illustrated. $2.00.

**Biology for Beginners.** By *Truman J. Moon*, Head of the Science Department, The High School, Middletown, N. Y., and *Paul B. Mann*, Head of the Department of Biology, Evander Childs High School, New York City. 1932. 741+xxxii pages, cloth. Henry Holt & Company.